

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

6606	Qy	TGGCGGCTCAGAGTACCGGAGGTGCGCAGCAGCGGTACATCACTACATATAACGAGC	6666	Qy	TCCACCTGATAACTTTGAAAGTCCCTGCCAACTACCTCTCTCCGAGTCTTTTCTCGG	7254	Qy	ATTACCAACCGGCACCTGTGCGGCTGTGTCTCCCTCTCTAGGAAAACCCGACGC	7314	Qy	CTCCCCAAGAGGCGCGGACAGTGGGCTTAAGTGGAGATCTCCATAGAGAGATGCCCTTC	7374	Qy	AACAGCTGGCCATTAAAGTCTTTGGCAGCCCCCCCCAAGCGCGGATTCAGGSCCTTCCA	7434	Qy	CGAGCTGGCCACCAAAAGTTTGGCAGCTCTCAACTCCGGCATTCAGGGCGCAATA	7494	Qy	AGAGAGTTCCATCTCTTCATGCCCCCTCGAGGGGAGCTTGGAGATCCAGACCTGG	7554	Qy	AGCCTGACAGGTAGAGCCCCAACCCCCCCCCCGAGGGGGGTGGCAGCTCCCGGCTCGG	7614	Qy	ACTCGGGTCTGGTCTACTTCTCCGAGGAGCAGATCCCGTCTGTGTCTCTCATGT	7674	Qy		7734	Qy		7794	Qy		7854	Qy		7914	Qy		7974	Qy		8034	Qy		8094	Qy		8154	Qy		8214	Qy		8274	Qy		8334	Qy		8394	Qy		8454	Qy		8514	Qy		8574	Qy		8634	Qy		8694	Qy		8754	Qy		8814	Qy		8874	Qy		8934	Qy		8994	Qy		9054	Qy		9114	Qy		9174	Qy		9234	Qy		9294	Qy		9354	Qy		9414	Qy		9474	Qy		9534	Qy		9594	Qy		9654	Qy		9714	Qy		9774	Qy		9834	Qy		9894	Qy		9954	Qy		10014	Qy		10074	Qy		10134	Qy		10194	Qy		10254	Qy		10314	Qy		10374	Qy		10434	Qy		10494	Qy		10554	Qy		10614	Qy		10674	Qy		10734	Qy		10794	Qy		10854	Qy		10914	Qy		10974	Qy		11034	Qy		11094	Qy		11154	Qy		11214	Qy		11274	Qy		11334	Qy		11394	Qy		11454	Qy		11514	Qy		11574	Qy		11634	Qy		11694	Qy		11754	Qy		11814	Qy		11874	Qy		11934	Qy		11994	Qy		12054	Qy		12114	Qy		12174	Qy		12234	Qy		12294	Qy		12354	Qy		12414	Qy		12474	Qy		12534	Qy		12594	Qy		12654	Qy		12714	Qy		12774	Qy		12834	Qy		12894	Qy		12954	Qy		13014	Qy		13074	Qy		13134	Qy		13194	Qy		13254	Qy		13314	Qy		13374	Qy		13434	Qy		13494	Qy		13554	Qy		13614	Qy		13674	Qy		13734	Qy		13794	Qy		13854	Qy		13914	Qy		13974	Qy		14034	Qy		14094	Qy		14154	Qy		14214	Qy		14274	Qy		14334	Qy		14394	Qy		14454	Qy		14514	Qy		14574	Qy		14634	Qy		14694	Qy		14754	Qy		14814	Qy		14874	Qy		14934	Qy		14994	Qy		15054	Qy		15114	Qy		15174	Qy		15234	Qy		15294	Qy		15354	Qy		15414	Qy		15474	Qy		15534	Qy		15594	Qy		15654	Qy		15714	Qy		15774	Qy		15834	Qy		15894	Qy		15954	Qy		16014	Qy		16074	Qy		16134	Qy		16194	Qy		16254	Qy		16314	Qy		16374	Qy		16434	Qy		16494	Qy		16554	Qy		16614	Qy		16674	Qy		16734	Qy		16794	Qy		16854	Qy		16914	Qy		16974	Qy		17034	Qy		17094	Qy		17154	Qy		17214	Qy		17274	Qy		17334	Qy		17394	Qy		17454	Qy		17514	Qy		17574	Qy		17634	Qy		17694	Qy		17754	Qy		17814	Qy		17874	Qy		17934	Qy		17994	Qy		18054	Qy		18114	Qy		181
------	----	---	------	----	---	------	----	--	------	----	---	------	----	---	------	----	---	------	----	---	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-------	----	--	-----

Qy	3306	TCCTGGGAGCGGACAGACTGCTGTGGGACATATTTTACACGAGCTTCCCGTGTCCGCCC	3365	Db	4375	TTCTGCCACTGTACCCCTCCGGGCTCCGTCACTGTGTCCATCTTAACATCGAGGAGG	4434
Db	3295	CGTGGGGGAGATACCGCCGCTGCGTGACATCATCAAGCGTTGCCCGTCTCTGCC	3354	Qy	4446	TGGCCCTTGGGAGGAGGAGATCCCTTCTATGGGAGGCGGATTCCTCCCTGTCTTACA	4505
Qy	3366	GACTTGTGTGGGAGGTCTCTTGTGGCCAGCTGATGGCTATATCTCAAGGGGTGAGTC	3425	Db	4435	TTGCTCTGTCCACACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCCCTCGAGGTGA	4494
Db	3355	GTAGGGCCAGGAGATATCTGTTGGGCCAGCCGAGGAAATGGTCTCCAAAGGGGTGAGGT	3414	Qy	4506	TCAAGGGAGGAGAGACATCTGTCTTGTGCATTTCAAAGAAAGTGTGACGAGCTCGCG	4565
Qy	3426	TTCTGCCGCCCATCACTGCTTACGCCAGCAGACAGTGGCTTTTGGGCACCATAGTGG	3485	Db	4495	TCAAGGGGAGAGACATCTCATCTTCTGCTCAAAAGAAAGTGTGACGAGCTCGCG	4554
Db	3415	TGCTGGCGCCCATCACGGCTGACGCCAGCAGACAGAGGCTCTTAGGGTGTATATCA	3474	Qy	4566	CGGCCCTTGGGGGTATGGGCTTGAACCTAGTGGCATCTACAGAGGTTGGAGCTCTCG	4625
Qy	3486	TGAGCATGAGGGGCGGACAGACAGACAGAGCTGGGGAATTTCAAGTCTCTGTCCACAG	3545	Db	4555	CGAAGCTGTGCGCATTTGGGCATCAATGCGGTGGCCCTACTACCGCGTCTTGAAGTGTCTG	4614
Db	3475	CCAGCTGACTGGCCGGGACAAAGAACCAAGTGAGGGGTGAGGTCCAGATCGTGTCAACTG	3534	Qy	4626	TATATCAAACTCAGGAGAGAGTAGTGTGTGCTGCCACCGACCGCCCTCATGACAGGGTATA	4685
Qy	3546	TCACTAGTCTTCTCGGAACATCATCTCGGGGTTTGTGGACTGTCTACATGGAG	3605	Db	4615	TATATCCGACACAGCGGATGTTGTGTGTGACCGATGCTCTCATGACTGGCTTTA	4674
Db	3535	CTACCCAAACCTTCTGSCAACGTGCATCAATGGGGTATGCTGGACTGTCTACACCGGG	3594	Qy	4686	CTGGGACTTTGACTCCGTGATCGACTGCAACGTAGCGGTCACTCAAGTTGTAGACTTCA	4745
Qy	3606	CTGGCAACAGACTCTGCGCGCTCACGGGCTCCGCTCACGAGATGTACTCCAGTGTCTG	3665	Db	4675	CCGGGACTTCGACTCTGTGATAGACTGCAACAGGTGTGTCACTCAGACAGTCAATTC	4734
Db	3595	CCGGAACGAGACCATCGCATCACCCAAAGGTCTCTCATCCAGATGTATACCAATGG	3654	Qy	4746	GTTTAGACCCCATTTCAACATTAACACAGATTTGCTCCTCAGAGGCTGTCTCAGCTA	4805
Qy	3666	AGGGGACTTAGTAGGGTGGCCAGCCCTCCGCTGAGCTAAATCTTTGGAGCCGTCACAGT	3725	Db	4735	GCCTTGACCTTACCTTTACCATTTGAGACCAACCGCTCCCGAGGATGCTGTCTCCAGGA	4794
Db	3655	ACCAAGACTTGTGGGCTGGCCGCTCTCAAGGTCCCGCTCATTTGACACCCCTGCACCT	3714	Qy	4806	GCCAGCGCGGGGTCCACGGGTAGGGAGAGCTGGGCATTTATAGTATGTTTCCACTG	4865
Qy	3726	GTGAGCGGTGACCTGTATCTGTGTACCGGGAACGTGTATGTATCATCCCGGTCTGAAGAC	3785	Db	4795	CTCAACCGCGGGGACAGACTGGCAGGGGAGCCAGGCATCTACAGATTTGTGGCACCGG	4854
Db	3715	GCGGCTCTCGGACTTTTACCTGTGTACGAGGACCGCGATGTCAITTCGGTGTGGCGGC	3774	Qy	4866	GTGAGCGAGCTCAGGAATGTTTGAAGTGTAGTGTCTGTGAGTGTCTACGACGAGGGG	4925
Qy	3786	GCGGGACAAACGGGAGCGCTACTCTCCCGAGACCTCTTCCACCTTGAAGGGTCTCT	3845	Db	4855	GGGAGCGCCCTCCGGCATGTTTCGACTCGTCCGTCTCTGTGAGTGTCTATGACGCGGCT	4914
Db	3775	GAGGTGATGAGCGGGTAGCTGTCTTTCGGCCCGGCCCAITTTCTACTTTGAAAGGCTCT	3834	Qy	4926	CCGCATGTGTATGAGCTCACACATCGGAGACACCGCTCAGGCTCAGGGGTATTTCAACA	4985
Qy	3846	CAGAGGCCCCGCTGCTATGCCAGGGGCGACCGCTGTGCGAGTCTTCCGGGCGAGCTGT	3905	Db	4915	GTCTTGTGTATGAGCTCAGCGCCGCGAGACTACAGTTAGGCTACGAGCGTATAGAAC	4974
Db	3835	CGGGGGTCCGCTGTGTGTGCCCCCGGGACACCGCGTGGGCTATTTCAGGGCCGCGGT	3894	Qy	4986	CGCCCGGTTTGGCTGTGTCGAAGACCATCTTGTAGTGTGTTGGGAGGAGTGTTCACCGGCC	5045
Qy	3906	GCTCTGGGGGCTGGCTAAGTTCATAGATTTTCATCCCGTGTGAGACATCTGACATCTGCA	3965	Db	4975	CCCGGGGCTTCCGCTGTGCCAGACCATCTTGAAATTTTGGGAGGGGCTTTTACCGGCC	5034
Db	3895	GCACCGGTGAGTGGCTAAGCGGTGAGCTTTATCCCTGTGGAGAACCTAGAGACACCA	3954	Qy	5046	TCAACACATPAGATGCCCACTTCTTTTCCAAACAAAGCAATCGGGGAAAATTTTCGCAT	5105
Qy	3966	CGCGTCCCGACCTTTAGTGACAAACAGCACACCTGCTGTGCCCGCCAGCCTATCAGG	4025	Db	5035	TCACTCATATAGATGCCCACTTCTATCCAGACAAAGCAGAGTGGGGGAACTTTCCTT	5094
Db	3955	TGAGATCCCGGCTTACGGAACAATCTCTCTCACCAAGCAGTGTCCCGAGCTTCCAGG	4014	Qy	5106	ACTTAAACAGCTACAGGCTACAGTGTGCGCTAGGGCCAAAGCCCGCCCGCTCTCTGGG	5165
Qy	4026	TGCGGTACTTGCATGCCCGGACTGGCAGTGGAAAGAGCACAAAGTTCTCTGTCTGCATATG	4085	Db	5095	ACCTGGTAGCGTACCAAGCCACCGTGTGGGCTCAGGGCTCAAGCCCTTCCCGCATCGTGGG	5154
Db	4015	TGGCCCACTGCTGCTCCACCGGACGCGTGAAGACCAAGGTCCTCGGCTCGGCTACG	4074	Qy	5166	ACGTGATGTGGAAGTGTGTTGACTCGACTCAAGCCCACTCTGTTGGGCCCCCAGACCTCTCC	5225
Qy	4086	CTGCTCAGGGGTATAAGTGTAGTGTATATTCCTCAGTGGCTGCGACCCCTGGGGTTG	4145	Db	5155	ACCAGATGTGGAAGTGTGATCCGCTTAAACCCACCTTCCATGGGCAACACCCCTG	5214
Db	4075	CAGCCAGGGCTACAAGGTGTGGTGTCTCAACCCCTCTGTGTCTGCAACCGTGGGGTTG	4134	Qy	5226	TGTACCGCTTGGGCTCTGTTTACCAACAGAGTCAACCTCTCACATCCCGTGGACGAAATACA	5285
Qy	4146	GGGCTACTTCTTAAGGCATGGCATCATTCACATTCAGACTTGGAGTCAAGGACTG	4205	Db	5215	TATACAGCTGGGGCTGTTCAGATGAAAGTCACTTACCGCACCAATCAACCAATACA	5274
Db	4135	GTGCTTACATGTCGAAGGCCCATGGGGTTGATCTTAATATCAGGACCGGGGTGAGAAC	4194	Qy	5286	TCGCCACTGTCATGCAAGCGGACTTGTAGGTCTATGACAGACATGGGTCTTGGCAGGGG	5345
Qy	4206	TGACGACCGGGCGCCATACGCTACTCCATATGGAATGGAATGCAATGCGGATGGGGCT	4265	Db	5275	TCAATGACATGATGTGGGCGGACCTGGAGGTCTGTACGAGCACCTGGGTCTCTGTGGG	5334
Db	4195	TTACACTGGCAGCCCATACGCTACTCCACTACGGCAAGTTCCTTGGCCGACGCGGGT	4254	Qy	5346	GAGTCTTGGCGCGCTGCGCGCTATTTGCTGGGACCGGGTGTGTTGATCATCTCGGCC	5405
Qy	4266	GTGCGGGCGGCTTACGACATCATATGTATGTATGAAATGCCATGCGGTGAGCTCTACCA	4325	Db	5335	GCCTTGGCTGTCTGTGGCGCGGTATTTGCTGTCAACAGGCTGCGTGGTCTATAGTGGGA	5394
Db	4255	GCTCAGGAGGTGCTTATGACATAATATTTGTGACGAGTGCACTCCACCGGATGCCCAT	4314	Qy	5406	GCTTGCACATTAACACGAGCGCTGTTTGGCGCGGACAGAGGAGTCTCTATGAGGCTT	5465
Qy	4326	CCATCTTGGCATGGAACTCTTGTATCAAGCAGACAGCTGGGTGAGCTCACTACTG	4385	Db	5395	GGATGCTGTCTGGGAGCGCGCAATTTATCTGACAGGAGGTCTCTTACCAGGAGT	5454
Db	4315	CCATCTTGGGCTCGGCACTGTCTTTGACCAAGCAGAGACTGCGGGGGCGAGATGTTGT	4374	Qy	5466	TTGATGAGATGGAGGAATGTGCTCTTAGGGCGGCTCTCATTTGAAGAGGGGACGGGATAG	5525
Qy	4386	TGCTGGCTACAGTACGCGCCCTGCGTCAGTGTGACAAACCCCGCCCAACATAGAGGAGG	4445				

Db 1087 GGAACGCAAACTCCCAACAGCAGCTTCGACGTATATCGATCTGCTTTGCGGAGCG 1146
Qy 1146 CCAGCTCTGCTCTGCCCTCTAGTGGGAGCCTCTGCGTGGGGTATGCTCGCAGCCC 1205
Db 1147 CCACCTCTGCTCGGCCCTCTAGTGGGAGCCTGTCGGGTCTGCTTTCTTTGTTGGTC 1206
Qy 1206 AAATGTTTCACTTCTGCGCAGCAGCAGCTGGTTGTCCAAGAGTCAATTTGCTCCATCT 1265
Db 1207 AACTGTTTCACTTCTCTCCAGCGCCACTGGAGCAGCAGCAAGACTCAATTTCTATCT 1266
Qy 1266 ACCCTGTATCATCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1325
Db 1267 ATCCGCGCCATATAACCGGTCATCGCATGCGATGGATATGATGAACTGTCGCCCTA 1326
Qy 1326 CGGCTACCATGATCTTGGCGTACCGATGCGTCCCGAGGTCATTTATAGACATCATTA 1385
Db 1327 CGGAGGTTGGTGTAGCTACGCTGCTCCGATCCCAAGCCATCATGACATGATCG 1386
Qy 1386 GCGGGCTCATTTGGGCGCTCATGTTGCGCTTGGCTTCTATGAGGGAGCGCTGGG 1445
Db 1387 CTGCTGCTACTGGGAGTCTTGGCGGCATAGGATTTCTCCATGGTGGGAACTGGG 1446
Qy 1446 CGAAAGTCTGTTGATCTTCTGTTGGCCCGCGGGTGGACGCGCAGCCATCTGTTG 1505
Db 1447 CGAAGTCTGTTGATCTGCTGCTATTTGCGCGCTGCGACGCGGAAACCCACGTCACCG 1506
Qy 1506 GGGGTTCTGCGCGCAGACCAACCGGCGCTCACAGCTTATTTGATGGGCGCCAGCG 1565
Db 1507 GGGGAAGTGGCGCGCAGCACCGGCTGGGCTTGTGGTCTCTTACACAGGCGCCAAAG 1566
Qy 1566 AGAAATCCAGCTCGTTAAACCAATGGCAGCTGGCACATCAACCGCAGCGCCCTGAACT 1625
Db 1567 AGAATCCAACTGATCAACCAACCAACCGCAGTGGCACATCAATAGCAGCGCTTGA 1626
Qy 1626 GCAATGACTCTTGGACACCGCTTTATCGGCTCTCTGTTCTATGTTTACACCCAGCTTCA 1685
Db 1627 GCAATGAAGCTTAAACCGGCTGTTAGCAGGCTCTTCTATCAGCAAAATTCAACT 1686
Qy 1686 GCTCAGATGTCGCGAGCAGTCTGCGCTGCGCAGTATCGAGGCTTCCGGTGGGAT 1745
Db 1687 CTTGAGCTGCTGAGAGTGGCAGCTGCGCAGCTTACCGATTTTGGCCAGGCT 1746
Qy 1746 GGGCGCTTTCGAATATGAGATAATGCTCAACCAATCCAGAGGATATGAGACCTTATTTGCT 1805
Db 1747 GGGCTCTATCAGTTATGCCAA-----CGGAAGCGGCTCGACGAAAGCCCTACTGCT 1800
Qy 1806 GGCACTACCAACAAAGCAGTGGGCTGCTGCTCGGAGAGTGTGTGGGCCAGTGT 1865
Db 1801 GGCACTACCCCTCAAGACCTTGTGGCATTTGCGCGCAAGAGCGTGTGTGGGCCGGTAT 1860
Qy 1866 ACTGTTTCAACCCAGCCAGTGTAGTGGGACGACGACAGGCTTGGAGCGCCACTT 1925
Db 1861 ATTGCTTCACTCCAGCCCGGTGGTGGGAGCGACGACAGGCTCGGCGCGCTTACCT 1920
Qy 1926 ACAGTGGGGAGATGAGACAGATGCTTCTCTATTGAACAGCACTCGACACAGCTGG 1985
Db 1921 ACAGCTGGGGTGAATATACGATGCTTCTGCTTAAACACACAGCGCACCGCTGG 1980
Qy 1986 GGTATGTTTGGCTGCACTGATGAATCTTTCTGGCTACACCAAGACTTGGCGGCAC 2045
Db 1981 GCAATTTGGTTCGGTTGTACTCGATGAACTCAACTGGATTCAACAAAGTGTGGAGCGC 2040
Qy 2046 CACCTGCGCTACTAGAGCTGATTTCAACGCGACGACGACCTTGTGTCGCCCGAGCT 2105
Db 2041 CCCCTTGTGTATCGGAGGGGTGGGCAACA-----CCTTGTCTCGCCCACTGATT 2094
Qy 2106 GTTTTGAAGAGCATCTGATACCACTTACTCAATGCGGCTCTGGGCGCTTGGCTCACGC 2165
Db 2095 GTTTCGCAAGCATCGGAAGCAGATCTCTCGGTGCGGCTCGGCTCGCTGGATTAC 2154
Qy 2166 CAAGGTGCTGATGATACCCCTACAGGCTCTGGCATTTACCCCTGCAAGTTAATATA 2225
Db 2155 CCAGGTGCTGCTGATACCCCTATAGGCTTTGGCACTATCTTGTACCATCAATTACA 2214

Qy 2226 CCATCTTTCAAAATAAGATGTATGTGGAGGGTGTGAGCACAGGCTCAAGCTGATGCA 2285
Db 2215 CCATATTTAAAGTACAGATGTACGTGGAGGGGTGAGCACAGGCTGGAAGCGGCTGCA 2274
Qy 2286 ATTTCACTCGTGGGATCGTTGCACTTGGAGGACAGACAGAGTCAACTGTCTCTCTT 2345
Db 2275 ACTGGAGCGGGCGGAACGCTGTGATCTGGAAGACAGGAGGTCGAGCTCAGCCAT 2334
Qy 2346 TGTGCACTCCAGCAGGAATGGGCAATTTTACCTTGTCTTCTTACTGCGACCTGCGCGCT 2405
Db 2335 TGTCTGTGTCACACAGAGTGGCAGGCTCTTCCGTGTTCTTTTACAGCCTGCGCAGCT 2394
Qy 2406 TGTGCACTGCTCTTCTCACTCCACCAAAACATCGTGACGCTGACAAATCATGTATGGCC 2465
Db 2395 TGTCCACCGGCTCATCCACCTCCACAGACATTTGGACGCTGAGTACTTGTACGGGG 2454
Qy 2466 TATCACTGCTCCCTCAAAAATACATCGTCCGATGGAGTGGGTAACTTATTTCTCTGC 2525
Db 2455 TAGGGTCAAGCATCGCTCTGGGCCATTAAGTGGGAGTACGTCTCTCTCTCTCTCTC 2514
Qy 2526 TCTTAGCGACCGCAGGTTTGGCGCTGCTTATGAGAGTCTCATCTTGTGGGCGCAGGCG 2585
Db 2515 TGTCTGACAGCGCGCGCTCTCTCTCTGCTGATGATGTTACTCATATCCCAAGCG 2574
Qy 2586 AAGCAGCACTAGAGAGTGTCTTGCACGCTGCGAGCGCAGTACTGCATGGCT 2645
Db 2575 AGCGGCTTTGGAGAACCTCGTAATCTCAATGCAAGCATCCCTGGCGGAGCGCAGCTC 2634
Qy 2646 TCTATATTTTCTCATCTTTTCTGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2705
Db 2635 TTGTGCTC 2694
Qy 2706 TAGCTACCTATTTC 2765
Db 2695 GAGCGCTCTACGCTTCTACGCGATGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 2754
Qy 2766 AACGCTTATGCTTATGAGCATCTGTCATGCGCAGTGGCAGTGGAGCGCTGCTGCTGTA 2825
Db 2755 AGCGGCTATACGCTATGAGCAGGAGTGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 2814
Qy 2826 TGATCACTCTCTTTACTCTCACCCCGGATATAAGACCTTCTCAGCGGCTTTTGTGTG 2885
Db 2815 GGTAAATGCGCTGACTCTGTCGCCATATTACAAGGCTACATCAGCTGCTCATGTGT 2874
Qy 2886 GGTGTGCTATCTCTGACCTTGGGGAGAGTATGGTCCAGAGTGGGACACCACTATGC 2945
Db 2875 GGTCTCAGTATTTCTGACAGAGTAGAAGCGCAACTGCACTGTGGGTTCCCGCTCTCA 2934
Qy 2946 AGGTGCGGCTGCGCTGATGCGATATGCGCGCTGCCATATTTACCCAGGCTGG 3005
Db 2935 AGTCCGGGGGGGGCGGATGCGCTACTTACTCATGTGTGTGTACACCGGCTGCTGG 2994
Qy 3006 TGTGTGACATTAACAAAGTGGCTTCTTGGCGGCTGCTTGGGCTTCTACCTCTTAAAGGTG 3065
Db 2995 TATTTGACATCAACAACTACTCTGCGCATCTTGGAGCCCTTTGGATTTCTTCAAGCCA 3054
Qy 3066 TTTTGAACGCGGTGCGCTGCTGCTGAGGCTCAGCTCTACTAGGATGTGCACTATGG 3125
Db 3055 GTTCTCTTAAAGTCCCTTACTCTGTCGCTTCAAGGCTTCTCCGATCTCGCGCTAG 3114
Qy 3126 CAAAGCATCTCGCGGGGAGGATGCTGCGATGCGGCTACTAGCCCTTGGCAGTGA 3185
Db 3115 CGCGAAGATAGCCGAGGCTATTAAGTGAATGGCCATCATCAAGTTAGGGGGCTTTA 3174
Qy 3186 CTGGCACTTATCATCATCACCACTCACCCCTTATGCGGATTTGGCTGCTAGTGGCTGC 3245
Db 3175 CTGGCACTTATGTTATTAACCATCTCACCCCTTCTCGAGACTGGGGCGCACACGCTGC 3234
Qy 3246 GGGACTCGCGGTGCGCTGAGCTTATCTTCAAGTCCGATGGAGAGAAAGTCAATGG 3305
Db 3235 GAGATCTGGCGCTGCTGGAACCAAGCTGCTCTTCTCCGAATGGAGCCAAAGCTCATCA 3294

QY	7194	TACCGGCTGGGACCGGCTGATTAAACCCACCGCTTGTGGAAATCGTGGAAAGCCAG	7253	QY	8274	TTCTCTGAAACATCGGCGGAAAGAGACCCCTATATGGTGTTCGTATGATACCCGAT	8333
DB	7195	TGCCCGTTTGGGCGGCGGAGTAAACCCCGCTAGTAGAGAGTGGAAAGCCCTG	7254	DB	8209	TCCTCGTGAAGCGTGGAGTCCAAAGAGACCCCGATGGGTTCTCGTATGATACCCGCT	8268
QY	7254	ATTACCAACCGGCACTGTGTGGGCTGTCTCCCTCTCTCTAGGAAACCCGACGC	7313	QY	8334	GCTTTGACTCAACCGTCACTAGAGAGACATCAGAGTCTAGAGTCCATATATCGGCT	8393
DB	7255	ACTACGAACACCTGTGTGCTCATGGCTGCGGCTACACCTCCAGGTCCTCTGTGC	7314	DB	8269	GTTTGTGACTCCACAGTCACTAGAGAGACATCCGTACGAGAGGCAATTTTCAATGTT	8328
QY	7314	CTCCCCAGGAGCGCGGACAGTGGGCTAGTGAAGACTCCATAGAGATGCCCTTC	7373	QY	8394	GCTCTCTGCGGAGAGGCGGACACTGCGCATACATCGCTACTAGAGACTTATCGTGG	8453
DB	7315	CTCCGCTTCGGAAGAGCGTACGGTGTCTCTACCGAATCAACCTATCTACTGCTGTG	7374	DB	8329	GTGACTGTGACCCCAAGCGCGCTCAAGTCCCTCACTAGAGGCTTTATGTTG	8388
QY	7374	AACAGTGTGGCATTAAGTCTTTGGGCGAGCCCCCAAGCGCGGATTCAGGCTTTCCA	7433	QY	8454	GAGGCTCTATGTTCAACAGAGGCGCAAACTCGCGGTACAGGCGTTCCCGCGCAGCG	8513
DB	7375	CCGAGCTTGCCACCAAAAGTTTGGGAGCTCTCACTTCCGCACTTACGGGCAATTA	7434	DB	8389	GGGCGCTCTTACCAATTCAGGGGGGAAACTGCGGTACCGAGGTGCGCGGAGCG	8448
QY	7434	CGGGGCGGCGCTGCGGATTCGCGCAGTCAGACGCTCTCTGATGATTTGGCCCTTCG	7493	QY	8514	GGGTGCTCACCACTTAGCATGAGGAAACCATCATCATGTCTAGTAAAGCTTTAGGGCT	8573
DB	7435	CGACAAACATCTCTG-----AGCCGCGCTTCTGGCTGCCCCCGACTCG	7482	DB	8449	GGTACTGACAACTAGTGTGTGTAACACCTCACTTGTCTACATCAAGGCGCGGAGCT	8508
QY	7494	AGACAGTTCCATCTCTTCCATGCCCCCTCGAGGGGAGCTTGGAGATCCAGACTGG	7553	QY	8574	GTAAAGCTCGAGGGATATCGCGCCCACTGCTGTATGCGCGATGATTTGTTGTA	8633
DB	7483	ACGTTGAGTCTTATCTTCCATGCCCCCTCGAGGGGAGCTTGGGATCCCGA-----	7537	DB	8509	GTGAGCGCGAGGGCTCCAGGACTGCACATGCTCGTGTGCGCAGCATTTAGTCTTA	8568
QY	7554	AGCCTGAGCAGTAGAGCCCCCAACCCCCCGGCGGCTGCGAGCTCCCGGCTCG	7613	QY	8634	TCTCAGAAAGCCAGGGGACCCGAGGAGGACGAGCGAACTGAGAGCTTTCAGGAGGCTA	8693
DB	7538	-----TCTCAGCGAGC	7548	DB	8569	TCGTGMAAGTCTCGGGGCTCCAGGAGGACGCGCGGCTGAGAGCTTTCAGGAGGCTA	8628
QY	7614	ACTCGGGCTCTGCTTACTTGTCTCGAGGAGGACGACTCCGCTGTGTCTCCATGT	7673	QY	8694	TGACAGGATTTCTCGCTCTCTGTGTGACCCCCCGAGACCGGAGTATGATTTGGAGCTGA	8753
DB	7549	GGTCATGTCGACGCTCAGTGTGTGGGCGGACACGGAAGATGTCTGTGTCTCAATGT	7608	DB	8629	TGACAGGATTTCTCGCTCTCTGTGTGACCCCCCGAGACCGGAGTATGATTTGGAGCTGA	8688
QY	7674	CATCTCTGGACGGGCTCTAATTAATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	7733	QY	8754	TAAACATCTGCTCTCAATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	8813
DB	7609	CTTATTTCTGGACGCGGCTCTCTACCCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	7668	DB	8689	TAAACATCTGCTCTCAATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	8748
QY	7734	TTAAACCTTTGACCACTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	7793	QY	8814	ACCTGACAGAGACCTTACCACTCAATTCGCGCGGCTGTGTGTGTGTGTGTGTGTGTGTGT	8873
DB	7669	TCAACGACTGAGCACTCTGT	7728	DB	8749	ACCTTACCGTGTGACCTTACCACTCAATTCGCGCGGCTGTGTGTGTGTGTGTGTGTGTGT	8808
QY	7794	AGAGCGCTCACTAAGGCTTAAAGGTAATTTGTATGAGTGAAGTGTGTGTGTGTGTGTGT	7853	QY	8874	CCCTGTCAATTCATGGTGGGAAACATCATCTAGTACGCGCGGCTGTGTGTGTGTGTGTGT	8933
DB	7729	GCACTGCTTGCAGGAGGAGAAAGTCACTTTCAGAGCTGCAAGTTCTGGAGCC	7788	DB	8809	CTCCAGTCAATTCCTGT	8868
QY	7854	ACTACGACTCAGTCTTAAAGGACATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	7913	QY	8934	TGCTCTGATGACACTTCTTCT	8993
DB	7789	ATTACGAGGACGTCTCAAGGAGTCAAGGAGGCTCAAGGAGGCTCAAGGAGGCTCAAGG	7848	DB	8869	TGAT	8928
QY	7914	TACCATGAGGAGGCTTGCAGTAAACCCCACTTCTGCAAGATCTAAATATGGGT	7973	QY	8994	TCAACTTTGAGATGTACGAGCGGTGTACTCTGCTGAGTCCCTTGGACCTCCCGAGCTATAA	9053
DB	7849	TATCGTAGGAGGCTTGCAGTAAACCCCACTTCTGCAAGATCTAAATATGGGT	7908	DB	8929	TTAACCTGTGAGATCTACGAGGCTCTACTCATAGAACCTGATCTACTCTCCATCA	8988
QY	7974	TTGGGCTAAGGAGTCCGAGCTTGTCTGGGAGGCGCTTAAACCATCAAGTCCGCTGT	8033	QY	9054	TTGAAAGGTTACATGGCTTGAAGCTTTTCTCTGACACATACACTCTCCCACTGA	9113
DB	7909	ATGGGCAAAAGAGCTCCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	7968	DB	8989	TTCAAAGAGTCTCATGGCTCAGCGCATTTTCACTCTCAAGTCTACTCTCCAGGTGAATCA	9048
QY	8034	GGAGAGCTCTGAGGAGTCAAGAACCAATTCGACCACTTATGGCCAAAGT	8093	QY	9114	CACGGTGGCTTACGCGCTCAGAAACTTGGGCGGCGGCTCAGAGCGTGGAGAGCC	9173
DB	7969	GGAAAGACCTTCTGGAAGACAGTGTAAACCAATAGACACTACCATCATGGGCAAGG	8028	DB	9049	ATAGGTGGCGCATGCTCTCAGAAACTTGGGCTTCCCGCTTGGAGCTTGGAGACAC	9108
QY	8094	AGGTGTTCTGCTGGAGCCCAACCAAGGGGCAAGAACGCTCGCTTATGTTTACC	8153	QY	9174	GGGACGCTGAGTCAAGGCGCT	9233
DB	8029	AGGTTTTCTGCTTACCTGAGAGGGGGTCTGTAAGCAGCTCTGCTCTCTCTCTCTCTCTCT	8088	DB	9109	GGGCGCGGAGCTCCGCTAGGCTTCTGTCTCAGAGGAGGAGGCTGCTGCTATATGGA	9168
QY	8154	CTGACCTCGGCTCAGGCTCTGAGAGAGATGCCCTTTATGACATTTACACAAACTTC	8213	QY	9234	GATATCTCTTCAATTTGGCGGTGAAGCAAGCTCAAACTCACTCTCAATTTGCCGGAAGCG	9293
DB	8089	CGACCTGGGCTGCGGCTGCGAGAGATGSCCTGTACGAGCTGTGTGTGTGTGTGTGTGTGT	8148	DB	9169	AGTACTCTTCACTGGGAGTGAAGCAAAAGCTCAAACTCACTCAATTAGCGCGGCTG	9228
QY	8214	CTCAGGCTGTGAGGGGCTCTTATGATTCAGTATTTCCCGCTCTCAGGGGTAGGT	8273	QY	9294	GCCTCTGATTTATCAGCTGT	9353
DB	8149	CCCTGCGCTGTGAGGAGCTCTACGATTTCACTCAACGAGGAGCGGGTTGAT	8208	DB	9229	CCCGCTGAGCTGTCTCGGTTGT	9288
				QY	9354	GGGTGCTGCTGCGGAGCCCGCTTATTTGCTTTTGGGCTACTCTCTACTTTTTTGTAGGG	9413

QY	2826	TGATCACTCTTTACTCTCACCCCGGGTATAAGACCTTCTCAGCCGGTGTGTTGGT	2885
Db	2815	GGTTAATGGCGCTGACTCTGCGCCATATTACAAAGCGCTACATCAGCTGTGATGGT	2874
QY	2886	GGTTGTGCTATCTTCTGACCCCTGGGGAAAGCTATGTCAGAGAGTGGGACCACTATGC	2945
Db	2875	GGCTTCAGTATTTTCTGACCAAGAGTAGAAGCGCAACTGTCAGCTGTGGTTCCTCCCTCA	2934
QY	2946	AGGTGCGCGTGGCGGTGATGGCATCATATGGGCGTTCGCGCATATTCTACCCAGGTGTG	3005
Db	2935	ACGTCCGGGGGGCGGATGCGGTCACTTACTCATGTGTGTGTACACCCGACTCTGG	2994
QY	3006	TGTTTGATATAACCAAGTGGCTCTTGGCGGTGCTTGGGCGTCTTACCTCTTAAAGGTG	3065
Db	2995	TATTTGACATCAACAACTACTCTGCGCATCTTCGACCCCTTTGGATTTCTAAGCCA	3054
QY	3066	CTTTGACGCGGTGCGGTATCTTGTGAGGCTCAAGCTCTACTGAGATGTGACCATGG	3125
Db	3055	GTTTGTCTTAAAGTCCCTACTTCTGCGCGTTCAAAGCCCTTCTCCGATCTGCGCGTAG	3114
QY	3126	CAAGGCATCTCGGGGGGCGAGTACGTTCAGATGGGCTACTAGCCCTTGGCAGGTGGA	3185
Db	3115	CGCGGAAGATAGCGGAGGTCTATGCTGCAAAATGGCCATCATCAAGTTTAGGGGCGCTTA	3174
QY	3186	CTGGCACTTACATCTATGACCACTCAACCCCTATGTCGGATTCGGCTGTAGTGGCTGC	3245
Db	3175	CTGGCACTATGTATTAACCATCTCAACCCCTCTTCGAGACTGGGCGCACACGGCTGC	3234
QY	3246	GGGACCTGGCGGTGCGGCTTGTAGCCTATCATCTTCAGTCCGATTCGAGAAAGTCAATTG	3305
Db	3235	GAGATCTGGCGGTGCTGGAACTAGTCTCTTCTCCGAAATGGAGACCAAGCTCATCA	3294
QY	3306	CTTGGGAGCGGAGACAGCTGCTTGTGGGACATTTTACAGGACTTCCGTGTCCGCC	3365
Db	3295	CGTGGGGGCGAGATACCGCGCGTGCCTGACATCATCAACGGCTTGCCTCTCTGCC	3354
QY	3366	GACTTGTGCGGAGGTCTCTCTTGGCCAGCTCATGCTATCTCCAAAGGGGTGGAGTC	3425
Db	3355	GTAGGGCCAGAGATACTGCTTGGGCGACCGACGGAATGCTCTCAAGGGGTGAGGT	3414
QY	3426	TTCTCGCCCTCATCTGCTTACCGCCAGCAGACAGTGGCTTGGGACCATAGTG	3485
Db	3415	TGCTGCGCCCATACCGCGTACCGCCAGCAGACGAGAGGCTCTTAGGGTGTATAATCA	3474
QY	3486	TGAGCATGACGGGGCGGACAAAGACAGAGACAGCTGGGAAATTCAGTCTCTGTCCACAG	3545
Db	3475	CCAGCCTGACTGGCCCGGACAAAACCAAGTGGAGGTGAGGTCCAGATCGTGTCAACTG	3534
QY	3546	TCACTCAGTCTTCTCGGAACATCCATCTCGGGGGTGTGAGACTGTCTACCATGGAG	3605
Db	3535	CTACCCAAACCTTCTGGCAACGTGATCAATGGGGTATGCTGAGACTGTCTACACGGGG	3594
QY	3606	CTGGCAACAGACTCTGGCGGCTCAAGGGTCCGGTCAAGCAGATGTACTCAGTGTG	3665
Db	3595	CCGGAACGAGGACCATCGCATCAACCAAGGCTCTGTCTATCCAGATGTATACCAATGTG	3654
QY	3666	AGGGGACTTGTAGGGTGGCCAGCCCTCTGGGACTTAAATCTTTGGAGCGGTGACCGT	3725
Db	3655	ACCAAGACCTTGTGGGTGGCGCGCTCTCAAGGTTCCCGCTCATTTGACACCCCTGACCT	3714
QY	3726	GTGGAGCGGTGACCTGTACTGTCAGCGGAACGCTGATGTATCCCGGCTCGAAGAC	3785
Db	3715	CGCGCTCTCGGACCTTTACTGCTGTCAGAGGACCGCGGATGTATTCCTGTCGCGCGG	3774
QY	3786	CGCGGACAAACCGGGAGCGCTACTCTCCCGGAGACTCTTTCCACTTTCAAGGGGTCTCT	3845
Db	3775	GAGGTGATAGCAGGGTAGCTCTTTTCCCGCGCCCATTTCTTACTTTGAAAGGTCTCT	3834
QY	3846	CAGGAGCCCGTGTATGCCCCAGGGGCGACGCTGTGAGAGTCTTCGGGAGAGTGTGT	3905
Db	3835	CGGGGGTTCGCTGTGTGCCCCCGGGACACCGCGTGGGCTTATTCAGGGCGCGGTGT	3894
QY	3906	GCTCTCGGGCGTGGCTAAGTCCATAGATTTTATCCCTGCTGAGACACTCGACATCGTCA	3965
Db	3895	GCACCGTGGAGTGGCTAAGCGGTGGAATTTATCCCTGTGGAGACTTAGAGACAACCA	3954
QY	3966	CGCGGTCCCGACCTTTTGTAGTGAACAAGCACCACTCTGCTGTGCCCCAGACCTATCAGG	4025
Db	3955	TGAGATCCCGGTGTTCAGGACAATCTCTCTCCACAGCAGTCCCGAGAGCTTCAGG	4014
QY	4026	TGCGGTACTTGTATGATCCCGACTGGCAGTGGAAAGAGCACCAAGTTCTGTGCGATATG	4085
Db	4015	TGGCCCACTGTATGCTCCACCGCAGCGGTAAAGAGCACCAAGGTCCCGGTGCGTAGC	4074
QY	4086	CTGCTCAGGGGTATAAGTGTAGTCTTAATCCCTCAGTGGCTGCACCTCGGGGTGTTG	4145
Db	4075	CAGCCAGGGCTACAAGGTGTGGTCTCAACCCCTCTGTGTGCAACGCTGGGCTTTG	4134
QY	4146	GGGCTACTTGTCTTAAGGCACATGGCATCAATCCCAATATAGGACTGAGAGTCAGACTG	4205
Db	4135	GTGCTTACATGTCCAAGGCCATGGGGTGTATCTAATATCAGGACCGGGTGAGAACAA	4194
QY	4206	TGACGACCGGGGGGCCATCAGTACTCCACATATGGCAATTTCTCGCGGATGGGGCT	4265
Db	4195	TTACCACTGGCAGCCCATCAGCTACTCCACCTACGGCAAGTTCCTTGCAGACGGCGGT	4254
QY	4266	GTGCGGGCGGCTACGACATCATATGTGATGAATGCCATGCCGTGGACTCTTACCA	4325
Db	4255	GCTCAGAGGTGCTTATGACATATAATTTGTGACAGTGCCTCCACGGATGCCACAT	4314
QY	4326	CCATCTTGGCATTCGGAAACAGTCTTGTATCAAGCAGAGACAGTGGGGTCAGACTAATG	4385
Db	4315	CCATCTTGGGCATTCGGCACTGCTTGTACAAAGCAGAGACTCGGGGGCAGACTGGTGTG	4374
QY	4386	TGCTGGCTACAGTACGGCCCTCGGTGAGTCAACACCCCAACCCCAACATAGAGGAG	4445
Db	4375	TGCTGGCCACTGCTACCCCTCGGGCTCGGTCACTGTGTGCCATCTTAAATCATCGAGGAG	4434
QY	4446	TGGCCCTTGGGCGAGGCGGAGATCCCTTCTATGGGAGGCGATTTCCTCTGTCTTACA	4505
Db	4435	TTGCTCTGTCCACCGGAGAGATCCCTTTTACGGCAGGCTATCCCTCTCGAGGTGA	4494
QY	4506	TCAAGGAGGAAAGACATCTGTATCTTGTCCATTTCAAAGAAAAAGTGTGACGAGCTCGCG	4565
Db	4495	TCAAGGGGGGAAAGACATCTCATCTTCTGCCACTCAAAGAAAGTGTGCGAGCTCGCGG	4554
QY	4566	CGGCTTTCGGGTATGGCTTGAATCAGTGGCATCTACAGAGGTTGGAGCTCTCG	4625
Db	4555	CGAAGCTGTGTCATTTGGGCATCAATGCGGTGGCTTACTACCGCGTCTTGAAGTGTCTG	4614
QY	4626	TAAATCCAACTCAGGAGACGTAGTGTGTCGCCAACCGACCGCTCTCATGACAGGGTATA	4685
Db	4615	TCATCCCGACAGCGCGGATGTTGTCTGTGTCGTCGACCGATGCTCTCATGACTGGCTTTA	4674
QY	4686	CTGGGACTTTGACTCCGCTGATGATGCAAGTGTAGCGGTCTCAAGTGTGTAGACTTCA	4745
Db	4675	CGGCGACTTTCGACTCTGTGATGAGTGTCAACACGCTGTGCTACTCAGACAGTCTGATTC	4734
QY	4746	GTTTACACCCCATTTTCACTTACCATTAACACAGATTTGTCCCTCAAGAGCTGTCTCAGTA	4805
Db	4735	GCCTTGGCCCTTACCTTTACCATTTAGACAACACCGCTCCCGGAGTGTGTCTCCAGGA	4794
QY	4806	GCAGCGCGGGGTGCGACGGGTAGGGAAGACTGGGCAATTTATAGGTATGTTTCCACTG	4865
Db	4795	CTCAACCGCGGGGAGGACTGGCAGGGGAGCGAGGATCTACAGATTTTGTGGACCGG	4854
QY	4866	GTGAGCAGCTCTCAGGAATGTTTGAAGTGTAGTGTCTGTGAGTGTCTACGACGCGGGG	4925
Db	4855	GGGAGCGCCCTCCGCGATGTTTCGACTCGTCCGCTCTCTGTGAGTGTATGACGCGGCT	4914
QY	4926	CGCATGTATGAGTCTACCATCGGAGACACCGTCAAGCTCAGGCTCAGGGGTATTTCAACA	4985
Db	4915	GTGCTTGTGTATGAGCTACGCGCGCGGAGACTACAGTTAGGCTACGAGGCTACATGAACA	4974
QY	4986	CGCGCGGTGTGCTGTGTGCCAAGACCATCTTGTAGTGTGAGGAGGAGTTCCTTTCACCGGC	5045

Copied from 09540843 on 05/19/2004

6655	TGACTACTGACAACTTTAAATGCGCGTGCAGATCCCATCGCCCGAAATTTTTCACAGAA	6714	7794	AGAGCGCTCACTAAGGCTAAAGGTAACCTTTTATAGGATGCAAGTCTCGACTCT	7853
6726	TGACGCGAGTGCAGATCCATAGCTTTGCGCCCGACACCGAGCGCTTTTTCGCGGATGAGG	6785	7729	GCAGTCTTGCACAAAGGAGAGAAAGTCACTTTTGACAGACTGCAAGTTCTTGACAGCG	7788
6715	TGACGCGGCTGCGCTTACATAGTCTTGCCTGCGCTTGCAGAGCTTCTGCTGCGGAGGAGG	6774	7854	ACTACGACTAGCTTTAAAGGACATTAAGTACGAGCGCTTCAAGGTACCGGACGCTCC	7913
6786	TCTCGTCTCGTGTGGCTTAAATTTATTTGCTGCGGTCCCGAGCTTCTTTGGAGCCCTG	6845	7789	ATTACGAGACCTGCTCAAGGAGGTCAAGAGCGCGCTCAAAAGGTGAAGGCTTAACCTTC	7848
6775	TATCATTTAGAGTAGGACTCCAGAGTACCGGTGGGTGCGCAATTAACCTTTGCGAGCCG	6834	7914	TCACCATGAGAGGCTTGCAGTTAAACCCCACTTCTGCAAGATCTAAATATGGGT	7973
6846	AACCGGACAGAGCTTATGATGCTGATGCTAAGATGCTATCATATCATATCAAGCGGAGAG	6905	7849	TATCCGTAGAGAGCTTGCAGCTGACGCGCCCACTTACGCCAATCAAGTTTGGGT	7908
6835	AACCGGACGAGCTGCTGAGCTTCACTGATGCTTCCATATTAACAGCAGAGG	6894	7974	TTGGGCTTAAGAGGTCCGAGCTTGTTCGCGGAGGCGCTTAAACCAATCAAGTCCGCTGT	8033
6906	CTGACGCGCGCTTTAGCGGGGTCAACCCCATCCGAGGCAAGCTCTCCGCGAGCC	6965	7909	ATGGGCAAAAGACGCTCGTTCGATGCGCAGAAAGCGCTAGCCACATCAACTCCGCTGT	7968
6895	CGGCGGGAGAGGTTGGGAGAGGCTCAACCCCTTCTATGGCAGCTCTCCGCGAGCC	6954	8034	GGAGAGCTCTCGAGGACTCAGAAACCAATTTCCCAACCAATTTGGCCCAAAATG	8093
6966	AGCTATCGGACCATCTGCGAGCCACCTGCAACCCACCGGCAAAAGCCTATGATGG	7025	7969	GGAAAGACCTTCTGGAAGACAGTGAACCAATAGACACTACCATCTGCGCCAAAGACG	8028
6955	AGCTGCTCGCTCCATCTCTCAAGCAACTTGCACCGCCCAACCATGACTCCCTCGAGCGG	7014	8094	AGGTGTTCTGCGTGGACCCCAACAGGGGGGCAAGAAACAGCTCCGCTTATCGTTTACC	8153
7026	ACATGCTGATGCTAACTCTCTGCGAGGAGGAGATGGGCGGCAACATCACCAGGCTTG	7074	8029	AGTCTTCTGCGTTCAGCTTGAGAGGGGGTCTGTAAGCAGCTCTCTCATCGTGTTC	8088
7015	AGCTCATAGAGGCTTAACCTCTGCGAGGAGGAGATGGGCGGCAACATCACCAGGCTTG	7074	8154	CTGACCTCGCGCTCAGGGTCTGCGAGAGATGGCCCTTTATGACATTTACACAAAACCTTC	8213
7074	AGTCTGGTCCAAAGTCTGCTGAGCTCTCTGACCCCAATGCTGCGAAGAGGAGCG	7133	8089	CCGACCTGGCGGTGCGGTGCGAGAGATGGCCCTGACGAGCTGTTAGCAAGCTCC	8148
7075	AGTCAGAGAAACAAAGTGGTGAATCTGAGCTCTCTGATCCGCTTGTGGCAGAGGAGTG	7134	8214	CTCAGCGGTGATGGGGCTTCTTATGATTCAGTATTTCCCGCTCAGCGGCTAGAGT	8273
7134	ACCTTGAGCTTCCATACCATCAGAAATACATGCTCCCGCAAGAGAGGTTCCCAACAGCTT	7193	8149	CCCTGGCGGTGATGGGAAGCTCTACGATTTCAATACCTCACAGGACAGCGGTTGAAT	8208
7135	AGCGGAGGCTCTGATCCCGCAGAAATCTGCGGAGTCTCGAGATTCGCGCGGCC	7194	8274	TTCTCTTGAAGCATGCGCGGAAAGAACGACCTATGCTTTTCTGATGATACCCGAT	8333
7194	TACCGGCTTGGGACGCGCTGATTAACACCGCTGCTGAGATCTGCGAAGAGGCGAG	7253	8209	TCCTGCTGAGAGCTGGAAGTCCAGAGACCCCGATGGGTTCTGTTGATGATACCGCT	8268
7195	TGCGGCTTGGGCGCGCGGAGCTACAAACCGCGCTAGTAGAGACGCTGGAAAGGCTG	7254	8334	GCTTTGACTCAACCGCTCAGTGAAGAGACATCAGGACTGAGGAGTCCATATATCGGCGCT	8393
7254	ATTACCAACCGGCTGCTGCGGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	7313	8269	GTCTTACTCCACAGTCACTGAGAGCGACATCGTACGAGGAGGCAATTTACCAATGTT	8328
7255	ACTACGAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	7314	8394	GCTCTTGGCGGAGGCGCCACACTGCGCATACACTGCTTACTGAGAGACTTTAGCTGG	8453
7314	CTCCCCCAAGGAGCGCGGACAGTGGGCTTAAAGTGAAGACTCCATAGGAGATGCCCTTC	7373	8329	GTGACCTGGAACCCCGCGCTGCGCTCATCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	8388
7315	CTCCGCTCGGAAAGAGTACGCTGCTCTCACCGAATCAACCTTACTTACTGCTTGG	7374	8454	GAGGGCTATGTTCAACAGCAAGGCGCAAACTGCGGCTACAGCGCTTGGCGCGCGAGCG	8513
7374	AACAGTGGCCATTAAGTCTTTGGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	7433	8389	GGGCGCTCTTACCAATTTAAGGGGGGAAACTGCGGCTTACCGAGGTCCGCGCGCGAGCG	8448
7375	CCGAGCTTGCCACCAAAAGTTTGGCAGCTCTCAACTTTCGCGCATTTACGGGCGACAATA	7434	8514	GGGTGCTCACTAGCTAGGAGGCAACCATCATCATGCTACGTAAGGAGCTTTAGCGGCTT	8573
7434	CGGGGCGGCGCTGCGGATTCGCGCAGTCAAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT	7493	8449	GCCTACTGACAACTAGCTGTGTTAAACCCCTCACTTGTCTACATCAAGGCGCGGCGAGCT	8508
7435	CGAACAACTCTCTG-----AGCCCGCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	7482	8574	GTAAGCTGAGGAGTAATCGCGCCCAATGCTGTTATGCGGCGATGACTTGGTGTCTCA	8633
7494	AGACAGGTTTCCATCTCTTCCATGCGCCCTCGAGGGGAGCTTGGAGATCCAGACCTGG	7553	8509	GTGAGCGCGGAGGCTCCAGGACTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	8568
7483	ACGTTGAGTCTATCTTCTCATGCGCCCTGCGGAGGAGCTGCGGAGTCCGA-----	7537	8634	TCTCAGAAAGCCAGGAGGAGCGGAGGAGCGGAGGAGCTTCAAGGAGCTTCAAGGAGGCTA	8693
7554	AGCCTGAGAGGAGTACAGCCCAACCCCGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	7613	8569	TCTGTGAAGAGTGGGGGCTCCAGGAGGAGCGCGGAGGCTTGAAGGCTTCAAGGAGGCTA	8628
7538	-----TCTCAGGAGCG	7548	8694	TGACCAAGTATCTGCGCTTCTGCTGAGCCCGCCCGAGACCGGAGTATGATCTGGAGCTGA	8753
7614	ACTCGGGTCTGCTTACTTCTGCTCGAGGAGGAGGAGTCTCGCTGCTGCTGCTGCTGCTGCTGCT	7673	8629	TGACCAAGTATCTGCGCTTCTGCTGAGCCCGCCCGGAGCCCGCCAGACCAAGATACGACTT	8688
7549	GGTATGCTGAGCGGCTGATGAGTGGGCGGAGACGAGAGATGCTGCTGCTGCTGCTGCTGCTGCT	7608	8754	TAAATCTGCTCTTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	8813
7674	CATATCTGAGCGGGGCTCTAATACTCTTGTAGTCCCGAAGAGGAGGAGGAGGAGGAGGAGG	7733	8689	TAAATCTGCTCTTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	8748
7609	CTTATCTGAGAGGCGACTCTGCTACCCCGTGGCTGCGAGAGACAAAGATGCGCA	7668	8814	ACCTGACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	8873
7734	TTAACCCCTTGAAGAACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	7793	8749	ACCTTACCGGTGAGCTTACCAACCCCTTACCAACCCCTTACCAACCCCTTACCAACCCCTTAC	8808

Qy	4506	TCAAGGAGGAGGACATCTGATCTTCTGCCATTCACAAAGAAAGTGTGACGAGCTCGCGG	4565	Db	5575	TTATCACCCCTGTCTCCAGACCAACTGCGAGAACTCGAGGTCTTCTCGGCGAGACACA	5634
Db	4495	TCAAGGCGGGAAGACATCTCATCTTCTGCCACTCACAAGAAAGTGTGACGAGCTCGCGG	4554	Qy	5646	TGTGGAACTTCATTAGCGGCATCCAACTCGCAGGACTATCAACACTGCCAGGAAACC	5705
Qy	4566	CGGCCCTTCGGGATATGGCTTGAATCTAGTGGGATCTACAGAGGTTGGAGCTCTCCG	4625	Db	5635	TGTGGAAATTCATCAGTGGGATACAACTATGCGGGGCTGTCAACGCTGCCTGTGTAACC	5694
Db	4555	CGAAGCTGGTGGCATCGGGATCAATCCGCTGGGCTACTACCGGGCTCTTGAAGCTGTCTG	4614	Qy	5706	CTGCACTAGCTTCCATGATGGCGTTTCAGTGGCGCCTCACCACTCGCTGTCAACAAGCA	5765
Qy	4626	TAATACCAACTCAGGGAGAGTGTGTGTGCTGCGCACCGAGCGCTCATGACAGGGGTATA	4685	Db	5695	CGGCCATGCTTCATTGATGGCTTTTACAGCTGCCGTCAACAGCCCACTAACCATCTGGCC	5754
Db	4615	TCATCCCGACACGCGGAGATGTTGTCTGTGTCGACCGATGCTCTCATGACTGGCTTTA	4674	Qy	5766	CCACTATCCTTCTCAACATTTTGGGGGCTGGCTAGCACTCCCAAAATTCACCAACCOCGG	5825
Qy	4686	CTGGGACCTTTGACTCCGCTGATCGACTGCAACGTAGCGGTCACTCAAGTTGTAGACTTCA	4745	Db	5755	AAACCCCTCTCTCAACATATTTGGGGGGTGGTGGCTGCCAGCTCGCGCCCGCGGTG	5814
Db	4675	CCGCGACTTCGACTCTGTGATGACTGCAACACGTGTGTCACTCAGACAGTGCATTTCA	4734	Qy	5826	GGGCCACTGGCTCTGTGTGCTGAGTGGGAGCTGCCGTAGGCACTATAGGCTTAG	5885
Qy	4746	GTTTAGACCCCACTTACCATATACACACAGATGTCCCTCAAGAGCTGTCTCAGGTA	4805	Db	5815	CCGCTACCGCCCTTTGCGGCGCTTAGCTGGGCGCCCATCGGACGCTTGGACTCG	5874
Db	4735	GCCTTGAACCTTACCTTTACCATTTGAGACAAACACGCTCCCGCAGGATGCTGTCTCCAGGA	4794	Qy	5886	GTAAGGTGTAGTGGACATCTCGGAGGGTATGTTGGCGGGCATTTTCGGGGCTCTCGTCCG	5945
Qy	4806	GCCAGCGCGGGTGCACAGGTAGGGGAAGACTGGGCATTTATAGTATGTTTCCACTG	4865	Db	5875	GGNAGTCTCTGTGGNACATTTCTTGAGGGTATGGCGGGCGTGGCGGAGCTCTTGTAG	5934
Db	4795	CTCAAGCGCGGGGACAGACTGGCAGGGGGAAGCCAGGCATCTACAGATTTGTGGACCGG	4854	Qy	5946	CAITCAAGATCATGTCTGGCGAGAACCTCTCATGGAGGATGTGCTCAACTTCTGCTCCTG	6005
Qy	4866	GTGAGCGAGCTTCAGGAATGTTTGACAGTGTAGTGTCTCTGTGAGTGTCTACGACGACGGG	4925	Db	5935	CATTCAAGATCATGAGCGGTGAGGTCCCTCTCCAGCGAGGACCTGGTCAATCTGCTGCCCG	5994
Db	4855	GGGAGCGCCCTCCGGCATGTTGAGCTCGTCCGTCTCTGTGAGTGTCTATGATGACGCGGCT	4914	Qy	6006	GAACTTGTCTCGGGTGTCTTGTAGTGGGAGTCACTGTGCGGGCCATCTGCGCGGAC	6065
Qy	4926	CCGATGTATGAGCTCACACATCGAGACACCGCTCAGGCTCAGGGCGTATTTCAACA	4985	Db	5995	CCATCTCTCGCTGAGGCCCTTGTAGTCTGTGCTGTGCGCAGCAATCTGCGCGCGC	6054
Db	4915	GTGCTTGGTATGAGCTCACGCCCGCGAGACTACAGTTAGGCTACGAGCGTACATGAACA	4974	Qy	6066	AGTGGGACCGGGGAGCGCGCTCCATGATGAATAGACTCATGCTCTTGTCTTCCCA	6125
Qy	4986	CGCCCGGTTGCTGTGTGCGACAGACCATCTTGAGTTTGGGAGGAGTCTTTCACCGGC	5045	Db	6055	AGTTGGCGCGGGGAGGGGCGAGTGTCAATGATGAACCGGCTAATAGCTCTCGCCTCCC	6114
Db	4975	CCCCGGGCTTCCGCTGTGCGAGACCATCTTGAAATTTTGGGAGGGCGTCTTACCGGGC	5034	Qy	6126	GAGGAATCACTCGCGCCCACTAGCTGACGAGGTGCGATGCTCGCAGCGGTGTGA	6185
Qy	5046	TCACACATAGATGCCCACTTCTTCCCAACAAAGCAATCGGGGAAATTTTCGCAT	5105	Db	6115	GGGGGAACCATGTTTCCCAACGCACTAGTCTCGCGGAGCGATGACGCGCGCGGTCA	6174
Db	5035	TCATCATATAGATGCCCACTTCTTCCCAACAAAGCAAGTGTGGGAGAACTTTCCTT	5094	Qy	6186	CCCACTACTTGGCTCTCTTACCAATACAGCGCTGTCTAGAGACTCTCAAACTGGATTA	6245
Qy	5106	ACTTAAAGCCTTACAGGCTACAGTGTGCTGAGGCGCAAGCCCGCCCGCTCTGCGG	5165	Db	6175	CTGCCATCTACAGCAGCTCTCACTTAACCCAGCTCTCTGAGGCGACTGCTCATGTGATAA	6234
Db	5095	ACCTGTGAGCGTACCAAGCAGCGTGTGCTGAGGCTCAAGCCCTCTCCCGCATCTGCGG	5154	Qy	6246	CTGAGACTGCCCACTCCCATCGCGCGCTCGTGTGCTCGCGATGTGTGGAGCTGGTTT	6305
Qy	5166	ACGTGTGTGGAAGTGTGACTCGACTCAAGCCCACTCGTGGGCGCCACACCTCTCC	5225	Db	6235	GCTCGAGGTGACCACTCCATGCTCGGTTCTGTGCTAAGGACATCTGGGACTGGATAT	6294
Db	5155	ACCGATGTGGAAGTGTGATCGCTTAAACCCCTTCAACCGGCGCAACACCCCTGCG	5214	Qy	6306	GCAACATCTTAAAGACTTTTAAATTTGGCTGACCTCCAAATTTATTTCCAAAGATGCCCG	6365
Qy	5226	TGTACCGCTTGGGCTCTGTACCAACAGAGGTCAACCTTCAACATCCCGTGACGAAATACA	5285	Db	6295	CGGAGTGTGAGCGACTTTAAGACTGGCTGAAAGCCCAAGCTCATGCCACAACCTGCTG	6354
Db	5215	TATACAGACTGGGCGCTGTTTCAAGATGAAGTCAACCTGACGACCCCAATCACCATAACA	5274	Qy	6366	GGCTCCCTTGTCTCTGTCAAAAGGGGTACAAGGGGTGTGGGCGGCGACCTGGCATCA	6425
Qy	5286	TGCGCACTGTGATGAAGCGGACTTGTGAGTCAATGACAGCAATGGGCTTGGCAGGGG	5345	Db	6355	GGATTCCTTGTCTGCTGCCAGCGCGGTATAGGGGGGTCTGCGGAGGAGACGGCATTA	6414
Db	5275	TCATGACATGATGTGCGGCGGACTGAGGCTCGTCAAGGACACTGGGCTGCTGTGGCG	5334	Qy	6426	TCACACACGGGTCTCTTGGCGGCAATATCTCTGGCAATCTCGCTTGGGCTCCATCA	6485
Qy	5346	GAGTCTTGGCGCGCTCGCGGCTTGTGCTGGCGACCGGCTGTGTGATCATCGGC	5405	Db	6415	TCACACCTGCTGCCACTGTGGAGCTGAGATCACTGGAATGTCAAAACGGGACGATGA	6474
Db	5335	CGGTCTGGCTGTCTGGCGCGTATTTGCTGTCAACAGGCTCGGTGTATAGTGGGCA	5394	Qy	6486	GAATCAAGGGGCTTAAGACTTCATGAATATCTGGCAGGGGACCTTTCTCTATCAATGTT	6545
Qy	5406	GCTTGCACATTAACACGAGCGCTGCTGTGGCGGACAGAGGCTCTCTATGAGGCTT	5465	Db	6475	GGATGCTCGTCTTAGGACTTCAGGACATGTGGAGTGGGAGCTTCCCAATTAAGCCT	6534
Db	5395	GGATGCTTGTTCGGGAAGCGGCAATTTATCTGACAGGGAGGTTCTCTACAGAGT	5454	Qy	6546	ACAGGAGGGCGAGTGGTGGCGGAAACCGCGCGCAAACTTTTAAAGTGTGCGCATCTGAGGG	6605
Qy	5466	TTGATGAGTGGAGGAATGCTCTTAGGGCGCTCTCATTTGAGAGGGGACGCGGATAG	5525	Db	6535	ACACCAAGGGGCGCTGTACTCTCCCTTCTGCTGCCGAACTATTAAGTTTGGCGGTGAGGG	6594
Db	5455	TCGATGAGTGAAGAGTGTCTCAGCACTTACCGTACATCGAGCAAGGATGATGCTCG	5514	Qy	6606	TCGCGGCTCTCAGAGTACGCGGAGTGAAGCAGCGGGTCAACCACTACATAAAGAGAC	6665
Qy	5526	CCGAGTGTCAAGTCCAGATCCAGGCTTATGCAAGCAAGCTTCCAAACAGCTCAAG	5585	Db	6595	TGTCTGAGAGGAATACGTGGAGATAAGCGGGTGGGAGCTTCCACTACGTATCGGTA	6654
Db	5515	CTGAGAGTTCAGAGCAGAGGCGCTCTGGCTCTCTGAGAGACCGGCTCCCGCAAGAGG	5574	Qy	6666	TCACCACTGATTAAGTGAAGTCCCTGCAACTACCTCTCCCGAGTTCTTTCTCGG	6725
Qy	5586	ACATACAAACCACTGTGAGGCTTTCATGGCGGCAAGGTAGAAACAAATTTCTGGGCAACACA	5645				

Db	2275	ACTGGACGGGGGCAACGCTGTGATCTGGAAGACAGGGACAGGTCCGAGCTCAGCCCAT	2334
Qy	2346	TGTTGACATCCACAGGAATGGCCATTTTACTTGTCTTACTCGGACCTCGCCGCT	2405
Db	2335	TGCTGTGTCCACACAGTGGAGGTCTTCCGTGTTCTTTACAGACCTGCCAGCT	2394
Qy	2406	TGTCGACTGTCTTCTCACCTCCACCAAAACATCGTGGACGTACAAATTCATGTGGCC	2465
Db	2395	TGTCACCGGCTCATCCACTCCACAGAACATTTGGAGCTGCAGTACTTGTGAGGG	2454
Qy	2466	TATCACCTGCTCCACAAAATACATCGTCCGATGGGAGTGGTAATCTTATTTCTGTC	2525
Db	2455	TAGGGTCAAGCATCGCTCTCGGCCATTAAGTGGAGTACGTCGTTCTCTGTTCTCTC	2514
Qy	2526	TCTTAGGGAGCGCAGGCTTTGGCTGCTTATGAGTGCATCTTGTGGGCCAGGCG	2585
Db	2515	TGCTTGAGAGCGCGCTCTGCTCTGCTTGTGGATGATGTTACTCATATCCAGCGG	2574
Qy	2586	AAGCAGCACTAGAGAGCTGGTCATCTTGACCGTGGAGCGAGCTAGCTGCAATGGCT	2645
Db	2575	AGCGGCTTTGGAGAACCTCGTAATCTCAATGCAGCATCCCTGGCCGGACGACGCTC	2634
Qy	2646	TCCTATATTTGTATCTTTTGTGCTGCTGTTGGTACATCAAGGTGCGGTAGTCCCT	2705
Db	2635	TTGTGCTCTTCTGCTGTTCTTCTGCTTTGCTGTTATCTGAAGGTAGGTGGTGGCCG	2694
Qy	2706	TAGTACCTATCCCTCACTGGCTGCTGCTTGTAGCTACTGCTCTAGCATTTGCC	2765
Db	2695	GAGCGGTCTAGCGCTTCTAGCGGATGCGCTCTCTCTGCTCTCTGCTGGCTGCTC	2754
Qy	2766	AACAGGCTTATGCTTATGAGCATCTGTGATGCGCAGATAGGAGCGCTCTGTGGTAA	2825
Db	2755	AGCGGCTACGCACTGCAACGAGGTGGCGGCTGCTGTGGCGGTTGTTCTTGTGTC	2814
Qy	2826	TGATCACTCTCTTACTCTCACCCCGGTATAGACCTTCTCAGCGGTTTGTGT	2885
Db	2815	GGTTAATGGCGCTGACTCTGTGCGCATATACAGCGCTACATCAGCTGTGTGCTGT	2874
Qy	2886	GGTTGTCTATCTCTACCTCGGGAGCTATGTCAGGAGTGGGACCACTATGC	2945
Db	2875	GGCTTCAGTATTTCTGACAGATAGAGCGCAACTGCGAGTGGGTTCCCGCCCTCA	2934
Qy	2946	AGGTGCGGCTGGCGGTGATGGCATCATATGGGCGCTCGCCATATTTCTACAGGTGG	3005
Db	2935	ACGTCCGGGGGGCGGATGCGCTCATCTTACTCATGTGTGTGTATACACCGACTCG	2994
Qy	3006	TGTTTGACATACCAAGTGGCTTTGGCGGTGCTGGGCTGCTTACTCTCTAAAGGTG	3065
Db	2995	TATTTGACATCACCAAACTACTCTGCGCATCTTTCGGACCCCTTTGGATTTCTTCAAG	3054
Qy	3066	CTTTGACGCGGTGCGGTACTTCTGAGGCTCAGCTCTACTGAGGATGTGACCATGG	3125
Db	3055	GTTTGTCTTAAGTCCCTTCTCTGCGGCTTCAAGGCTTCTCCGATCTGCGCGTAG	3114
Qy	3126	CAAGGCATCTCGCGGGGCGAGGTACGTCCAGATGGGCTACTAGCCCTTTGGCAGTGA	3185
Db	3115	CGCGAAGATAGCGGAGTCAATAGTGCATAATGGCCATCATCAAGTTAGGGGCGTTA	3174
Qy	3186	CTGGCACTTACATATGACCACTCTACCCCTATGTGCGATGGGCTGCTAGTGGCTGC	3245
Db	3175	CTGGCACCTATGTATAAACCATCTCACCCCTCTTCGAGAGTGGGCGCAACAACGCGCT	3234
Qy	3246	GGGACCTGGGCTGCGGCTTGTAGGCTATCATCTTCACTGCTCCGATGGAGAAAGTCAATG	3305
Db	3235	GAGATCTGGCGGTGGCTGTGGAAACAGTCTCTTCTCCGATGGAGAACCAAGTCTATCA	3294
Qy	3306	TCTGGGAGCGGAGACAGCTCTTGTGGGACATTTTACAGGACTTCCCGTGTCCGCC	3365
Db	3295	CGTGGGGGCGATACCGCGCGTGGGTGACATCATCAACGGCTTGGCGCTCTCTGCC	3354
Qy	3366	GACTTGTGGAGGTCTCTCTGGCCCGAGCTGATGCTATPACTCTCAAGGGGTGGAGTC	3425
Db	3355	GTAGGGGCCAGGAGATCTGCTTGGGCGAGCGGAAATGGTCTTCAAGGGGTGGAGGT	3414
Qy	3426	TTCTCGCCCCCATCTGCTTACCGCCAGCAGACAGTGGCCCTTTTGGGCACCATAGTG	3485
Db	3415	TGCTGGCGCCCATCAGCGGTACGCCACAGCAGAGAGGCTCTCTAGGCTGTATAATCA	3474
Qy	3486	TGAGCATAGCAGGGGCGGACAGACAGAAACAGCTGGGAAATTCAGGTCTCTGTCCACAG	3545
Db	3475	CCAGCCTGACTGCGCGGGACAAAACCAAGTGGAGGTGAGGTCCAGATCTGTGTCACTG	3534
Qy	3546	TCACTCAGTCTCTCTCGGAACATCCATCTCGGGGGTTTTGGAGCTGTCTACCATGGAG	3605
Db	3535	CTACCCAAACCTTCTCTGGGACGTGATCAATGGGGTATGCTGGACTGTCTACACGGGG	3594
Qy	3606	CTGGCAACAGACTCTTGGCGGCTCACGGGGTCCGCTCAGCGAGATGTACTCCAGTGTCTG	3665
Db	3595	CCGGAACGAGGACCATCGCATCACCCCAAGGCTCTCTCATCCAGATGTATACCAATGTG	3654
Qy	3666	AGGGGACCTTAGTAGGGTGGCCAGCCCTCGGGACTAAATCTTTTGGAGCGGTGACAGT	3725
Db	3655	ACCAAGACCTTGTGGGCTGGCCGCTCTCAAGGTTCCCGCTCATTTGACACCTCGACCT	3714
Qy	3726	GTGGAGCGGTGACCTGTACTGTCACGCGAAGCTGATGTATCCCGGCTCGAAGAC	3785
Db	3715	GGCGCTCTCGGACCTTTTACCTGTCAGGAGCAGCGCATGTCATCCGTCGCGCGG	3774
Qy	3786	GCGGGACAAACGCGGAGCGCTACTCTCCCGGAGACTCTTTCCACCTTTGAAGGGTCTCT	3845
Db	3775	GAGGTGATAGCAGGGGTAGCTGCTTTCCGCCCGCCCATTTCTTACTTTGAAGGCTCTCT	3834
Qy	3846	CAGGAGCGCGGTGCTATGCCCCAGGGGCGACGCTGTGCGAGTCTTCCGGGCGAGCTGTGT	3905
Db	3835	CGGGGGTCCGCTGTTGTGCGCGCGGACACGCCGTGGGCTTATTTCAAGGGCGCGGTGT	3894
Qy	3906	GCTCTCGGGCGGTGCTTAACTCATAGTTTCAATCCCGTTGAGACACTCGACATCGTCA	3965
Db	3895	GCACCGTGGAGTGGTAAAGGGGTGAGCTTTATCCCTGTGGAGAACCTTAGAGAACCA	3954
Qy	3966	CGCGGTCCCCCACCCTTTAGTGACAAACAGCAGCACCACTGCTGTGCGCCGAGACTATCAGG	4025
Db	3955	TGAGATCCCGGTGTTTACCGGACAACTCTCTCCACAGCAGTCCCGAGCTTCCAGG	4014
Qy	4026	TCGGGTACTTGTGATGCCCGGACTGGCAGTGGAGAGACCAAGTCTCTGTGCGCATATG	4085
Db	4015	TGGCCACCTGCTATGCTCCACCGGAGCGGTAAAGAGCACCAAGGTCCCGGCTGCGTAGC	4074
Qy	4086	CTGCTCAGGGGTATAAGTGTAGTCTTAATCCCTCAGTGGCTGCCACCTCGGGTTTG	4145
Db	4075	CAGCCAGGGGTACAAGGTGTGGTGTCTAACCCCTCTGTGCTGCAACGCTGGGCTTTG	4134
Qy	4146	GGCGGTACTTGTCTAAGGACATGGCATCAATCCCAACATTAGGACTGGAGTCAGGACTG	4205
Db	4135	GTCTTTACATGTCCAAGGCCATGGGTGTGATCTTAATATCAGGACCGGGGTGAGAACAA	4194
Qy	4206	TGACGACCGGGGCGCCATCAGTACTCCACATATGGCAATTCCTCGCGATGGGGCT	4265
Db	4195	TTACCACTGCGACCCCATCAGTACTCCACCTACCGCAAGTTCCTTTGCGCGCGGGT	4254
Qy	4266	GTGCGGCGGCGCTTACGACATCATATGTGATGAATGCCATGCCGTGGACTCTTACCA	4325
Db	4255	GCTCAGAGGTGCTTATGACATAATTTGTGACGAGTGCCACTCCACGGATGCCACAT	4314
Qy	4326	CCATCTCTTGCATCGGAAACAGTCTTGTATCAAGCAGAGACAGCTGGGGTTCAGACTTAACTG	4385
Db	4315	CCATCTTGGGCTCGGACATGCTCTTGAACCAAGCAGAGACTGCGGGGCGAGACTGGTGTG	4374
Qy	4386	TGCTGGCTACAGCTACGCCCCCTGGGTGAGTGAACACCCCGACCCCAACATAGAGAGG	4445
Db	4375	TGCTCGCACTGCTACCCCTCCGGGCTCCGCTCACTGTGTCCCATCTCTAACTCGAGAGG	4434
Qy	4446	TGSCCTTGGCAGGAGGCGGAGATCCCTTCTATGGAGGGGATTCCTCTCTTACA	4505
Db	4435	TTGCTGTCTCCACCAACCGGAGAGATCCCTTTTACCGCAAGGCTATCCCTCTCGAGTGA	4494

Qy 126 TCCCGGAGAGCCATAGTGGTCTCGGAAACCGGTGAGTACACCGGAATTCGCCGGAGAC 185
Db 127 TCCCGGAGAGCCATAGTGGTCTCGGAAACCGGTGAGTACACCGGAATTCGCCAGGAC 186
Qy 186 TGGGTCTCTTCTGGATAAACCCACTCTATGCCCGGCCATTTGGCGTGCCTCCCGGAG 245
Db 187 CGGTCTCTTCTGGATAAACCCCGCTCAATGCTCGSAGATTTGGCGTGCCTCCCGGAG 246
Qy 246 CTGTAGCCGAGTAGCGTTGGTGTGGAAGGCCCTGTGTGTAATTCGCTGATAGGTGCTT 305
Db 247 CTGTAGCCGAGTAGTGTGGTGTGGAAGGCCCTGTGTGTAATTCGCTGATAGGTGCTT 306
Qy 306 GCGAGTCCCGGGAGTCTCGTAGACCGTGCACCATGAGCACAAATTCCTAAACCTCAA 365
Db 307 GCGAGTCCCGGGAGTCTCGTAGACCGTGCACCATGAGCACAAATTCCTAAACCTCAA 366
Qy 366 GAAAAACCAAGAAACCAACCAACCGTGCACCATGAGCACAAATTCCTAAACCTCAA 425
Db 367 GAAAAACCAAGAAACCAACCAACCGTGCACCATGAGCACAAATTCCTAAACCTCAA 426
Qy 426 AGATCGTTGGCGAGTACTTGTTCGCGCGCAGGGGCCCGAGTTGGGTGTGCGCGCA 485
Db 427 AGATCGTTGGCGAGTACTTGTTCGCGCGCAGGGGCCCGAGTTGGGTGTGCGCGCA 486
Qy 486 CAAGGAAGACTTCGGAGCGGTCCAGACCAAGTGGAGCGCCAGCCCATCCCTAAAGATC 545
Db 487 CGAGGAAGACTTCGGAGCGGTCCAGACCAAGTGGAGTAGACGTAGAGCTATCCCAAGGAC 546
Qy 546 GCGCTCCACTGGCAATTCCTGGGGAACCAAGGATACCCCTGGCCCTATACGGGATG 605
Db 547 GTCGGCCGAGGCGAGACCTGGGCTCAGCCCGGGTACCCCTGGCCCTCTATGGCAATG 606
Qy 606 AGGAGCTCGCTGGCGAGGATGCTCTGCTCCCGAGGTCCCGTCCCTCTTGGGGCC 665
Db 607 AGGAGCTCGCTGGCGAGGATGCTCTGCTCCCGAGGTCCCGTCCCTCTTGGGGCC 666
Qy 666 CCAATGACCCCGGCATAGTTCGGCAACAGTGGGTAAAGTCAATGATACCTTAACGTGG 725
Db 667 CCACAGACCCCGGCATAGTTCGGCAACAGTGGGTAAAGTCAATGATACCTTAACGTGG 726
Qy 726 GCTTTGCGGACCTCATGGGTACATCCCTGTGTGGCGCCCGCTCGCGGGGTGCGCA 785
Db 727 GCTTCGCGGACCTCATGGGTACATCCCTGTGTGGCGCCCGCTCTTGGAGGCGCTGCCA 786
Qy 786 GAGCTCTCGCGCATGGGTGAGAGTCTGAGAGACGGGTAAATTTGCAACAGGGAAT 845
Db 787 GGGCCCTGGCGCATGGGTGAGAGTCTGAGAGACGGGTAAATTTGCAACAGGGAAT 846
Qy 846 TACCCTGCTGCTCTTTCTATCTTCTGTGTCGCTGTGCTGTGCTGATCACACCCCGG 905
Db 847 TTTCTGCTGCTCTTTCTATCTTCTGTGTCGCTGTGCTGTGCTGATCACACCCCGG 906
Qy 906 TCTCGCTGCGGAGTGAAGAACATAGTACCGGCTACATGGTGAATACGATGACATGCA 965
Db 907 CTTGAGCTTACCAAGTGGCAATTTCTCGGGGCTTTACCATGTGACCAATGATGGCCTA 966
Qy 966 ATGACAGATTAATCTGCGAGCTCCAGCTGCTGCTCTCCAGTCCCGGGTGGTCCCGT 1025
Db 967 ACTGAGTATGTGTACGAGCGGCGATGCCATCTCTGACATCGCGGGGTGTGCTCTT 1026
Qy 1026 GCGAGAAAGTGGGGAATGATCTCAGTGTGGATACCGGTCTCACGAATGTGGCGGTG 1085
Db 1027 GCGTTCGAGAGGTAAAGCTCGAGGTGTGGGTGGGGTGACCCCAACCGTGGCCACCA 1086
Qy 1086 AGCGGCCCGGCGCTCAGCGAGGCTTGGGAGCGACATCGACATGATGTTGATGTCG 1145
Db 1087 GGGACGGCAAACTCCCAACACGAGCTTCGAGGTATATCGATCTGCTGTGCGGAGCG 1146
Qy 1146 CCAAGTCTGCTGCGCTCTAGTGGGAGCTCTGCGGTGGGGTGTGCTCGAGCC 1205
Db 1147 CCAAGTCTGCTGCGCTCTAGTGGGAGCTCTGCGGTGGGGTGTGCTCTGTTGTTGGT 1206
Qy 1206 AAATGTTTCTGCTCGCGCAGCACCACTGGTTTGTGCAAGACTGCAATTTGCTCATCT 1265

Db 1207 AACTGTTTACCTTCTCTCCAGCGCCTAGGACGCAAGACTGCAATTTGTTCTATCT 1266
Qy 1266 ACCCTGATACCTGACATCGGACACCGCATGGACATGGACATGATGATGATGATGATG 1325
Db 1267 ATCCCGGCCATTAACGGGTCTATCGCATGGCATGGATGATGATGATGATGATGATG 1326
Qy 1326 CGGTACCATGATCTTTGGGTGATCGCATGCTGTCCTCCGAGGTCTATATAGACATCATTA 1385
Db 1327 CGGACGCTTGGTGTAGCTCAGCTGCTCGGATCCCAACGCCATCATGACATGATCG 1386
Qy 1386 GCGGGCTCATTTGGGCGTATGTTGCGCTTGGCTTACTTCTCTATGACAGGAGCGTGG 1445
Db 1387 CTGCTGCTCAGTGGGAGTCTTGGCGGATAGAGCTATTTCTCTCATGTTGGGAACTGG 1446
Qy 1446 CGAAGTCTGTTGTCATCTTCTGTTGGCGCGCGGGGTGACCGCGCACCATCATGTTG 1505
Db 1447 CGAAGTCTGTTGTTGCTGCTATTTGCGGGGTGACCGGGAACCCACGTCAACG 1506
Qy 1506 GGGTTCTCGCGCAGACACCGGCGCTCACAGCTTATTTGACATGGGCCCCAGGC 1565
Db 1507 GGGGAAGTCCCGCGCACACGCTGGCTTGTGGTCTCTTTACACAGGCGCCAGC 1566
Qy 1566 AGAAATCCAGTCTGTTAAACCAATGGCAGCTGGCACATCAACCGCACCGCCTGAAT 1625
Db 1567 AGACATCACTGATCAACCAACCGGAGTTGGCACATCAATAGCACGGCTTGAAT 1626
Qy 1626 GCATGACTCTTGCACACCGGCTTATCGCTCTCTGTTCTACACCGACAGCTTCACT 1685
Db 1627 GCAATGAAGACCTTAAACCGGCTGTTAGCAGGGCTCTTCTATCAGCAAAATTTCACT 1686
Qy 1686 CGTCAGATGTCGGAACGATGTCGCTCGCGAGTATCAGGCTTCCGCGTGGAT 1745
Db 1687 CTTGAGCTGCTGAGAGTTGGCAGCTGCGAGCTTACCGATTTTGGCAGGCT 1746
Qy 1746 GGGCGCTTGAATATGAGGATAATGTCAACCAATCAGAGGATATGAGACCTTATGCT 1805
Db 1747 GGGTCTTATAGTTATGCCAA-----CGGAAGCGGCTCGACGAACGCCCTACTGCT 1800
Qy 1806 GGCATACCAACAGGAGTGTGGGTGCTCTCGGAGAGACTGTGTGGCCAGTGT 1865
Db 1801 GGCATACCTCAAGACCTTGGGATTTGTCGCGCAAGAGCGTGTGGCCCGGTAT 1860
Qy 1866 ACTGTTTCAACCCCGAGCTGAGTGGGACACGACAGGCTTGGAGCGCCACTT 1925
Db 1861 ATTGCTTCACTCCAGCCCGTGTGTGGGAAACGACCGACAGTCTGGGCGGCTTACT 1920
Qy 1926 ACACTGGGGAGAAATGACAGATGTTCTTCTATTTGAACAGCACTCGACCACTG 1985
Db 1921 ACAGCTGGGTGCAAAATGATCGGATGCTCTGCTTAAACCAACAGGCGCACCGCTG 1980
Qy 1986 GGTATGTTTGGCTGACGCTGAGTGAATCTTCTGCTGCTACACCAAGACTTGGGGCG 2045
Db 1981 GCAATTTGGTTCGTTGTACTCGATGAACTCAACTGGATTCCAAAAGTGTGGGAGCG 2040
Qy 2046 CACCTGCGCTACTAGAGTGAATCAACCGCAGCAGGACCTGTTGTCGCCCGAGCT 2105
Db 2041 CCCCCTGTTGTCGAGGGGTGGGCAACAA-----CCTTGTCTCTGCTCCCTGAT 2094
Qy 2106 GTTTAGGAAGCATCTGATACACTTACTCTCAAAATGGCGCTCTGGGCGCTTGGCT 2165
Db 2095 GTTTCGCAAGCATCGGAAGCCACATCTCTCGGTGGGCTCCGCTCCCTGATTAAC 2154
Qy 2166 CAAAGTGCCTGATCGACTACCTTACCTGAGCTTACCCCTGACAGTGAATTAATA 2225
Db 2155 CCAGGTGATGTTGCTGACTACCGCTATAGGCTTTGGCACTATCTTGTATCAATTA 2214
Qy 2226 CCATCTTCAAAATAAGGATGATGAGGAGGCTTGGACACAGGCTCACGGTGCATGA 2285
Db 2215 CCATATTTCAAGTCAAGATGATGAGGAGGCTTGGACACAGGCTTGGAGGCGCTG 2274
Qy 2286 ATTTCACTGCTGGGATCGTTGCAACTTTGGAGGACAGAGAGTCAAGTGTCTCTT 2345

Copied from 09540843 on 05/19/2004

Qy	3842	TCCTCAGGAGCCGGTCTATGCCAGGGGCCAGCTCTCGAGTCTTCCGGCAGCT	3901	Qy	4922	GGGGCCGCTGATGAGCTCACACCATCGGAGACACCGTCTAGGCTCAGGGGTATTTTC	4981
Db	3843	UCAUCCGAGGGCCCGUGUCUGUCWAGGGGACACCCCGUGGCUUUGUACAGAGGCC	3902	Db	4923	GGGGCAGCCUGUACGAGCUUACACCUUGUGAGACUACGUGAGACUCCGGGCUUUAUC	4982
Qy	3902	GTGTGCTCTCGGGCGTGGCTAAGTCCATAGATTTTCATCCCGCTGAGACATCGACATC	3961	Qy	4982	AACACGCGCGTGTGCTGTCGCAAGACCATCTTTGAGTTTGGGAGGAGTTTTCACC	5041
Db	3903	GUGUGCCAGGGGUGUAGCAAAUUAUUGACUUAUCCCGUGGAUACACUGAURUC	3962	Db	4983	AACACGCGCGUUGCCGUAUGUACAGACCAACUGAGUUCUGGAGGAGCGGCUUACA	5042
Qy	3962	GTACGCGGTCCTCCACCTTTAGTGACACAGACACACCTGCTGTCGCCAGACCTAT	4021	Qy	5042	GGCTCACACATAGATGCCATCTTCTTCCAAACAAAGCAATCGGGGAAAAATTTTC	5101
Db	3963	GCCACAGGACGCCAGUUCUCUGAACACAGURCGCGCAGCUGUGCCCAAGUCUAC	4022	Db	5043	GGUCUCAC	5102
Qy	4022	CAGTGGGTACTTTGATGCTGCGGCTGAGTGGGAGGAGGAGGAGGAGGAGGAGG	4081	Qy	5102	GCATCTTAACAGCTTACAGGCTACAGTGTGCGCTAGGCGCAAAAGCCCCCGCTTC	5161
Db	4023	CAGUGGCUUACUUGACAGGACCAACAGGACGCGGAGGAGGAGGAGGAGGAGG	4082	Db	5103	GCRUAUUAAGGCGCUACAGGCGCACAGUAGGCGCAGGCGCAAAAGCCCCCUUCG	5162
Qy	4082	TATGCTGCTCAGGGGTATAAGTCTAGTCTTAATCCCTCAGTGGGTGCCACCTGGGG	4141	Qy	5162	TGGGAGCTGATGGAAGTGTGACTCGACTCAAGCCACACACTCGTGGGCGCCACACT	5221
Db	4083	UAUGCCAGUCAGGGGUAUUAAGUACUCGUACUAAUCCUCUGUGCGGCGCACAUUGU	4142	Db	5163	UGGAGCGUGAGUGGAGUGUUAACUAGGCGUCAAACCUACACUGACUGGUCACCC	5222
Qy	4142	TTTGGGCGTACTTGTCTAAGGACATGCGATCAATCCCAACATAGGACTGGATCAGG	4201	Qy	5222	CTCTGTGACCGCTTGGCTCTGTTCACACGAGTCACTCCCTCACATCCCGTGAAGAA	5281
Db	4143	UUUGGGCCUACAUUGCCUAAAGCCACCGGAUACCCUAAUACAGAACTGGAGUGCGG	4202	Db	5223	CUCUUGUACCGUUGGGGCGGUGACCAUAGAGUACCUUGACGACACCCCGUGAGAA	5282
Qy	4202	ACTGTGACGACCGGGCGCCATCAGCTTCTCAGTCTTCCATATGCGAAATTCCTCGCGATGGG	4261	Qy	5282	TACATCGCCACCTGTCATGCAAGCCGACTTGTAGGTCTATGACACGACATGGGTCTTGCA	5341
Db	4203	ACCUUUAACACCGGGGACUCUUAUACUUAUCCUUAUUGCAUUAUUGCAGAGGUA	4262	Db	5283	UACAUCCGACGUGCAUGCAAGCUGACUUGAGUACUAGCAAGCUCUAGUGGUCUGCG	5342
Qy	4262	GGTGTGCGGGCGCGCTACGACATCATATGATGATGATGATGATGATGATGATGATG	4321	Qy	5342	GGGGGAGTCTTGGCGCGCTGCGCGGTATTTGCTGGGACCGGGTGTGTTTGCATCATC	5401
Db	4263	GGCUGGACCGGCGGCUAUGACUACUUAUGCGACGAAUGCAUUCAGUGGACGCU	4322	Db	5343	GGGGGGGUGUAGCCCGGUGGAGUACUGGCGGCGACUGGCGUUAUUCUACUUAU	5402
Qy	4322	ACCAACATCTTGGCATCGGAACAGTCTTGTATCAAGCAGACAGCTGGGTGAGACTA	4381	Qy	5402	GGCGGCTTGCACATTAACAGCGAGCGCTGCTGTCGCGCAAGAGAGGTCTCTATGAG	5461
Db	4323	ACUACCAUCCUUGGCAUUGGAAACAGUCUCUUGACCAAGCUGAGACCGCAGGCGU	4382	Db	5403	GGCGGCGUACACCUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAG	5462
Qy	4382	ACTGTGCTGCTACAGCTAGCCCTCGGCTGAGTGAACCCCGACCCCAACATAGAG	4441	Qy	5462	GCTTTTGTAGATGAGATGAGGAAATGTGCTCTAGGGCGGCTCTCATTTGAGAGGGG	5521
Db	4383	GUGUYUUGGCGACAGCCGCUCCGUGUACUUGGUGUACUCCUACUAGUACUAGAG	4442	Db	5463	GCCUUGAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAG	5522
Qy	4442	GAGTGGCTCTTGGGCGAGGCGAGATCCCTTTCTATGGAGGGCGATTCCTCTGCT	4501	Qy	5522	ATAGCGGAGATGCTGAAGTCCAAAGGCTTATTTGACGACGAGCTTCCAAACAGCT	5581
Db	4443	GAGGUGCCUUGGUCACGAGGCGAGUCCUUAUUGGCAAGCUUAUCCCUAGCU	4502	Db	5523	AUGCGGAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAG	5582
Qy	4502	TACATCAAGGAGGAGACATCTGATCTTCTGCGCATTCAGAGAAAGTGTGACGCTC	4561	Qy	5582	CAAGACATCAACCCACTGTGAGGCTTCAATGCGCCCAAGGTAGAACAAATTTCTGGG	5641
Db	4503	UUAUCAAAGGGGCGAGACUUAUUGUCCAUUUAUCAAAGAAAGUAGUAGUAGUAG	4562	Db	5583	CAAGCAURCAGCAGCUACUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAG	5642
Qy	4562	GGGCGGCTTGGGCTGAGGCTTGAACCTGAGTGCCTGAGTGCCTGAGTGCCTGAG	4621	Qy	5642	CACATGTGGAATCTTATTTAGCGGCTTCAATACCTCGCAGGACTTATCAACATGCGC	5701
Db	4563	GCAGGCGCCUCCGGGCGAUGGUGUACUUGCCGUGUACUUAUAGGUGUUCGACUC	4622	Db	5643	CACAUGGAGAACUUAUACAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAG	5702
Qy	4622	TCCGTAATCAACCTCAGGAGAGCTAGTGTGCTGCGCACCGACCGCTCATGACAGG	4681	Qy	5702	AACCTGTGAGTGTCTCATGATGCGGCTTCAAGTGCCTCAGCTCAGCTCAGCTCAG	5761
Db	4623	UCCGUUAUACCAUUAUAGGAGAGGUGGUGUUGGUGUAGUAGUAGUAGUAGUAGU	4682	Db	5703	AUUCUGCAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAG	5762
Qy	4682	TATCTGGGACTTTGATCTCGTGTGATCGATGCGAAAGTGTGAGTGTGAGTGTGAG	4741	Qy	5762	AGCACCATCTCTCTTCAACATTTTGGGGGCTGCTGCTAGCTCCCAATTTGACACACC	5821
Db	4683	UACACCGGCGACUUGACUCUGUACUAGUAGUAGUAGUAGUAGUAGUAGUAGUAG	4742	Db	5763	AGCACCAUCCUUAUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAG	5822
Qy	4742	TTCACTTTAGACCCCATTCACCATTAACACACAGATGTCCTCAGAGCGTGTCTCA	4801	Qy	5822	GGCGGGGCGACTGCTTGTGTGCTGAGTGGGCTGAGTGGGCTGAGTGGGCTGAGTGG	5881
Db	4743	UUCAGCCUAGACCCCAUCCUACCAUACCAUACCAUACCAUACCAUACCAUACCAU	4802	Db	5823	GCCGAGGACACYGGUUCUUGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAG	5882
Qy	4802	CGTAGCAGCGCGGGTGCACGGGTAGGGGAGAGCTGGGCAATTTATAGTATGTTTC	4861	Qy	5882	TTAGTGAAGTGTAGTGTGACATCTCTGCGGAGGCTATGCTGGGGCTCTC	5941
Db	4803	CGUAGUACUAGAGGAGAGACUUGGAGGGGCGGAGUUGGCGUUAACAGGAUAGUUG	4862	Db	5883	CUGGUAAGAUACUGGUGGACGUGUUGGCGCGGUAUGGCGCAGGCAUUAUAGGGG	5942
Qy	4862	ACTGTGAGCGAGCTCAGGAATGTTTGAAGTGTGCTGTGAGTGTGAGTGTGAGGCA	4921	Qy	5942	GTCCATTCAGATCATGTCTGGGAGAGCCCTCCATGAGGAGTGTGCTCACTCTGCTG	6001
Db	4863	UCAGYGARRGGCGGUGUUGGAGUUGUUGACAGCGUAGUGUCYUGGCGAGUAGUAG	4922	Db	6002	GUAGCUUUAAGAUCAUGAGCGGAGAGCCCAACCGUAGAGAGAGAGAGAGAGAGAG	6002
				Qy	6002	CCTGGAATTTCTCTCCGGGTGCTTGGTGTGAGTGGGAGTCTATCTGCGGCGGCTTCTG	6061

Db 1623 A A U G C A A U G A C A G C Y U A S A G A C G G G U U U M U G C U U C Y U G K U U A C M M C R C A R G U U C 1682
 Qy 1682 A A C T C G T C A G A T G T C C C G A A C C A T G T C G C G T G C G C A G A T A T C G A G G C C T T C G G G T G 1741
 Db 1683 A A C A G C U C U G C G C C C G A G C G C U G U U C U G C G C G G G U G A C G A Y U U C G A U C 1742
 Qy 1742 G G A T G G G G C C C T T G C A A T A T G A G G A T A T G T C A C C A A T C A G A G A T A T G A C C C T A T 1801
 Db 1743 G G C U G G G A C C U U G A A U A C G A A C C A A C G U C A C C A A C G A U G R G G A C A U A G C G C G U A C 1802
 Qy 1802 T G C T G G C A C T A C C A C A A G C A G T G G G T G T C T C C G A A G A C T G T G T G T G G C C C A 1861
 Db 1803 U G C U G C A U A C C C C G A G C C U G C G G C A U G C U C C G C G C U A G G A C G G U U U C G G A C C G 1862
 Qy 1862 G T G T A C T G T T T C A C C C C C A G C C A G T G T A G T G G G C A C C A C C A C C A G C C T T G A G C G C C C 1921
 Db 1863 G U C U A U G Y U U C A C C C C A G C C C U G U G C G G G C A C C A C U G A C A G C A G G G C G U A C C C 1922
 Qy 1922 A C T T A C A C G T G G G G A G A A T G A G A C A G A T G T C T C T A T T G A A C A G C A C T C G A C C A C C G 1981
 Db 1923 A C C U A C A C C U G G G R G N A A C G A G A C G A U G U C U C C U G R A A A U A G C A A G A C C C C C G 1982
 Qy 1982 C T G G G T C A T G T T C G G T C C A C G T G A A C T C T T T G G T A C C A A G A C T T G C G G C 2041
 Db 1983 C G A G G A C U U G G U C G C U C A C Y U G A U A C C G G A C U G G U C A C U A A G A C A U G C G G U 2042
 Qy 2042 G C A C C A C C T G C G T A C T A G A C T G A C T T C A A C G C C A G C A C G A C C T G T T G T G C C C A C G 2101
 Db 2043 G C A C C A C C U U G C G C A U A G G A A G A C A C A A C A G C A C U C U G A U U A U U G U C C C C A C A 2102
 Qy 2102 G A C T G T T T A G A A G A T C C T G A T A C A C A T T A C T C A A T G C G C T C T G G C C C T G G C T C 2161
 Db 2103 G A C U U U A G A A G A C C A C C A G A U C A C U A C U A U A G U G A G A G A G C C C U U G U A 2162
 Qy 2162 A C G C A A G T G C T G A T C G A C T A C C C C T C A G C C T C T G C A T T A C C C T G C A C A G T T A A C 2221
 Db 2163 A C U C C A G G U C C U G U A G A C U A C C C U A U A G R Y U G U G C A U U A U C C G U C A C U G U A A C 2222
 Qy 2222 T A T A C C A T C T T C A A A T A G A T G T A T G T G G A G G G T T G A G C A C A G C C T G C A G G T G C A 2281
 Db 2223 U U C A C A C A U C U Y A A G C G C G A U A U A G A G A G G G U G A G A U C A U C U C C G C A G C A 2282
 Qy 2282 T G C A A T T T A C T C G T G G G A T C G T G C A C T T G G A G C A G A C A G A G A G T C A A C T G T C T 2341
 Db 2283 U G C A A C U U A C G C G G A U C C U G C A G A C U G A A G A U G G A U A G G G Y C A G C A G A U 2342
 Qy 2342 C C T T T G T G C A T C C A C C A G A A T G G C C A T T T A C C T T G C T T A C T C G G A C C T G C C C 2401
 Db 2343 C C A C U G C A U U C C A C A C U A G A G G C G G U Y U C C A U G C U C U C U G A C C U A C C A 2402
 Qy 2402 G C C T T G C A C T G T C T T C C A C C T C C A C C A A A A C A T G T G A C G T A C A A T T C A T G T A T 2461
 Db 2403 G C A C U A U C C A C U G C C C U A U G C A C C C A C C A A A A C A U C G U G A C C U G C A G A C C U Y A C 2462
 Qy 2462 G G C C T A C C T G C C C T C A A A A T A C A T C G T C C G A T G G A G G G T A A T A C T C T T A T T C 2521
 Db 2463 G G A C U U C C G C G C U G A C A G A U A C U G G A A G U G G A G U G G A U C C U C C U U U U C 2522
 Qy 2522 C T G C T C T T A G C A C C C A G G G T T G C C C T G T A T G A T G T C A T C T T T T T G G G C C A G 2581
 Db 2523 U U G U G U G C A C A C C C A G R U C U G A U G C U U U G A U G C U A C A U A C U G G C C A A 2582
 Qy 2582 G C C A A G C A C T A G A A G C T G T G T C A T T T G C A C G C T G C G A G C C A G T A G C A A T 2641
 Db 2583 G C C A A G C G C G C U G A G A G C U A C A C U U G C A C U C C G C U A G Y C U G C U A G U G C A A U 2642
 Qy 2642 G G C T C T A T A T T T G T C A T T T T T C G T G C T G T T G T A C A T A A G S T G G G T A G T C 2701
 Db 2643 G G U C C G U G U G U U U C A U C U U U A C A G C G C C U G G U A C U U A A A G G G C A G G G U G U C 2702
 Qy 2702 C C C T A G C A C C T A T T C C C T C A C T G G C C T G T G T C C T T A G C C T A C T G C T C C T A G C A T T G 2761
 Db 2703 C C C U G G C C A C G U A C U C U G U B U C G G C U U R U G U G U C C U C C U C C U C C U A G U C C U G G C Y U A 2762
 Qy 2762 C C C A A C A G C T T A T G T C T T A T G A C C A T C T G T C A T G C C A G A T A G A G C G C T C T G C T G 2821
 Db 2763 C A C A G C A G C U U A U G C C U U G G A C G C U C U G A A C A A G G G A A C U G G G C U G G C C A U A U A 2822
 Qy 2822 G T A A T G A T C A C T C T C T T T A C T C T C A C C C C C G G T A A G A C C C T T C T C A G C C G T T T T T G 2881
 Db 2823 G U A U A U A U C C A U C U U A C U A C C C C A G C A U A C A A G A U C C U C C U G A G C C G U C A G U G 2882
 Qy 2882 T G T G T T G T G T A T C T T C T G A C C C T G G G G A A G C A T A T G T T C A G A G T G G G C A C A C C T 2941
 Db 2883 U G G U C U C U C C U A C A U G C U G U C C U G C C A G G C C C A G A U C A G C A A U G G G U C C C C C 2942
 Qy 2942 A T G C A G T G C G G G T G C T G A T G C A T C A T A T G G C G T G C G C A T A T T C T T A C C A C C G T 3001
 Db 2943 C U G A G G U C C G A G G G G C G U A C G G G A U C A U C U G G G U G C U G A U C U A C A C C C A C C G C 3002
 Qy 3002 G T G G T T T T G A C A A A C A A G T G G C T T T G G G C G T C T T G G G C C T C T T A C C T C C T A A A A 3061
 Db 3003 C U U G U U U G A G U C A C A G A A U G U U A G C A A U C C U G G G C C U G C U A C C U C C U A R A 3062
 Qy 3062 G T G C T T T G A C G C G G T G C G T A C T T C T C A G G G T C A C G C T C T A C T A G A G A T G T C A C C 3121
 Db 3063 G C G U C U C U A C C G A U A C C G U A C U U U G U G A G G C C A C G C U U G C A G A G U G U A C C 3122
 Qy 3122 A T G G C A A G C A T C T C G C G G G G C A G T A C T C C A G A T G C C T A C T A G C C C T T G C A G G 3181
 Db 3123 C U G G U A A A C A C C U G C G R G G G C U A G U A C U C A G A U C G U U R A U C A C C A U A G G C A G A 3182
 Qy 3182 T G G A C T G G C A C T T A C A T T A T G A C C A C C T C A C C C T A T G T C G A T T G G C T G T A G T G C 3241
 Db 3183 U G A C C G C A C U A C A U A C A C C A C C U C C U C C C U U A C A A C U U G G C G C C C A G G U 3242
 Qy 3242 C T G C G G A C C T C G C G G T C G C C T T A C T C T T C A G T C C A T C T T C A G T C C A T G A G A A A G T C 3301
 Db 3243 U U R C G G A C C U G G C A A U C C C U G A G C C U G U G U C A G C C A A U G A G A A G A G U C 3302
 Qy 3302 A T T G T C G G G A C G G A G A C A G T C T T G T G G G A C A T T T A C A C G A C T T C C C T G T C C 3361
 Db 3303 A U U G U G U G G G G C U G A C A G U G C G U G U G A G A C A U C C U G A U G C C C G C G U C C 3362
 Qy 3362 G C C C A C T T G T C G G A G T C C T C T T G C C C A C A G T G A T G G T A T A C T C C A A G G G T G G 3421
 Db 3363 G C G A G C U A G U A G G A R G U C U C U G C G C C C U G C C G A C C U A C A C C U C C A A G G G U G 3422
 Qy 3422 A G T C T T C G C C C A T C A C T C T T A C G C C C A G C A G A C A G T G G C C T T T T G G C A C A T A 3481
 Db 3423 A A K C U C U A G C U C C A U A U A C U C U A C A C U C A G C A A A C U C U G U G U C C U C G G U G C U A U C 3482
 Qy 3482 G T G T G A G C A T A C G G G G C C A C A G A C A G A C A G C T G G G A A T T C A G T C C T G T C C 3541
 Db 3483 G U G U C A G C U A A C G G C C G C A C A A A A U A G A G C A G G C U G G C A G S U C C A G S U C U G U C C 3542
 Qy 3542 A C A G T C A C T C A G T C C T C T C T C G A A C A T C C A T C T C G G G A C T T T T G T G A C T G T C T A C C A T 3601
 Db 3543 U C G U C A C A A A C C U U U U G G G A C A U C A U U U C G G G C U C C U C U G A C A G A U A U A C 3602
 Qy 3602 G G A G C T G G C A A C A A G A C T C T G C C C G C T C A C G G G G T C C G G T C A C C A G A T G T A C T C A G T 3661
 Db 3603 G G G C U G G U A U A A G A C C U U G C C C C C C A A G G A C C A G U C A C U C A G A U A C A C C A G C 3662
 Qy 3662 G C T A G G G G A C T T A G T A G T G G C C C A G C C C C C C T G G A C T A A A T C T T T G A G C C G T G C 3721
 Db 3663 G C A A A G G G G A C C U C U G G G A U G G C C C C C C G G A C U A A G A C A U A G A C C C C U G 3722
 Qy 3722 A G T G T G A G C G G T G A C C T G A C C T G G T C A G C C G A A C C G T A G T C A T C C G G T C C A 3781
 Db 3723 A C C U G G G G C C U A G A C C U C U A C C U G U C A C C C A A A C C U G A U G A U C U C G G U C C G 3782
 Qy 3782 A G A C G C G G G A C A A C C G G G A G C T A C T C T C C C A G A C C T C T T T C A C C T T G A G G G 3841
 Db 3783 A G A A A A G A U C C G A C G G G U G A U A C U C U C G C C A A G G C C C C C U C A C C C C A A A G A 3842

COUNTRY: US
 ZIP: 20036
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/925,695
 FILING DATE: 19920807
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 287402/91
 FILING DATE: 09-AUG-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 360441/91
 FILING DATE: 05-DEC-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Weillacher, Robert G.
 REGISTRATION NUMBER: 20,531
 REFERENCE/DOCKET NUMBER: 06/87-48009
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 659-2811
 TELEFAX: (202) 659-1462
 TELETYPE: WUI 64470
 INFORMATION FOR SEQ ID NO: 6:
 TYPE: NUCLEIC ACID
 STRANDEDNESS: single
 TOPOLOGY: linear

US-07-925-695-6

Query Match 62.18; Score 6028.6; DB 1; Length 9511;

Best Local Similarity 59.8; Pred. No. 0;

Matches 5687; Conservative 1673; Mismatches 2149; Indels 0; Gaps 0;

Qy	2	CCCCCCCCCTAATAGGGCGACACCTCCCGCATGAATCACTCCCTGTGAGGAACTACTGTGTC	61
Db	3	CGCGCCCGGUGGGGGGACACUCCCGCAUGAUAUCCUCCUGUGAGGAAUUAUCCUGUC	62
Qy	62	TTACGCGAGAAAGCGTTAGCGATGGCGTTAGTATGATGTCGTACAGCTTCCAGGCCCC	121
Db	63	UUCACGCGAGAAAGCGUAGCGUAGCGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGU	122
Qy	122	CCCCCTCCCGGAGAGCATAGTGTCTCGGAAACCGGTGAGTACACCGGAAATTCGCGGGA	181
Db	123	CCCCCTCCCGGAGAGCGCAUAGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG	182
Qy	182	AGACTGGGTCTCTTTGGATAAACCCTATCTATGCGCGGCGCATTTGGGCGTGCCTCCGCG	241
Db	183	AGACUGGGUCCUUCUUGGAUAAACCAUUCUAGUGUGUGUGUGUGUGUGUGUGUGUGUGUG	242
Qy	242	AAGACTGCTAGCCGAGTAGCGTTGGTTGCGAAAGGCGTTGTGTACTGTGCTGATAGGTT	301
Db	243	AAGACUGUAGCCGAGUAGCGUUGGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG	302
Qy	302	GCTTGGAGTGCCTCCGGAGTCTCGTAGACCGTGCACATGAGCAGAAATCTTAAACCT	361
Db	303	RCUUGCGAGUGCCCCGGGAGGUCUGUAGACCGUGCAUAGGAGCAUAGGAGCAUAGGAGG	362
Qy	362	CAAGAGAAACCAAAAGAAACCAACCGTGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG	421
Db	363	CAAGAGAAACCAAAAGAAACCAACCGTGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG	422
Qy	422	GGCGAGATCGTTGGCGGAGTATCTTTGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	481
Db	423	GGUCAGAUCCUGCGGAGUUAUUAUUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG	482
Qy	482	CGGACAGGAAGACTTTCGGAGCGGTCCGAGCGCATGTTGGAAGGCGGCGGCGGCGGCGGCGG	541
Db	483	CGGACAGGAAGACTTTCGGAGCGGTCCGAGCGCATGTTGGAAGGCGGCGGCGGCGGCGGCGG	542

Qy	542	GATCGCGCTCCACTGCGCAAAATCCTGGGGAACACAGAGATACCCCTTGGGCGGCTATACGGG	601
Db	543	GAUCGGCGCTCCACTGCGCAAAATCCTGGGGAACACAGAGATACCCCTTGGGCGGCTATACGGG	602
Qy	602	AATGAGGAGTCCGCTGGGCGAGATGGCTCCTGTCCCGCGGAGTTCCGCTCCCTCTTGG	661
Db	603	AACGAGGGUUGGGGCGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG	662
Qy	662	GGCCCCAATGACCCCCCGGCATAGTCTCGGCAACGCTGGGTAAAGTTCATCGATACCTTAACG	721
Db	663	GGCCCCAATGACCCCCCGGCATAGTCTCGGCAACGCTGGGTAAAGTTCATCGATACCTTAACG	722
Qy	722	TGCGGCTTTGCGGACCTCATGGGTATACCTCTGTGTGGGCGGCGGCGGCGGCGGCGGCGG	781
Db	723	UGUGUUGUUGCGGACCCUAGGGGUAUACUCCUUGUGUGGUGGUGGUGGUGGUGGUGGUGG	782
Qy	782	GCCAGAGCTCTCGGCGATGCGGTGAGAGTCTTGGAGAGCGGGGTAAATTTGCAACAGGG	841
Db	783	GCCAGAGCTCTCGGCGATGCGGTGAGAGTCTTGGAGAGCGGGGTAAATTTGCAACAGGG	842
Qy	842	AACCTAACCGGTTGCTCTCTTTCTATCTCTTCTGCTGCGCCCTGCTGCTGCTGCTGCTGCTG	901
Db	843	AAUUAACCGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG	902
Qy	902	CCGCTCTCCGCTGCGGAAAGTGAAGATCATGATGCGGCTATCATGTTGATGATGATGATG	961
Db	903	CCAGUGUGCGAGUGGAAAGTGAAGATCATGATGCGGCTATCATGTTGATGATGATGATG	962
Qy	962	ACCAATGACAGCATTAACCTGCGGAGCTCCAGGCTGCTGCTCCACGCTCCCGGCGGCGG	1021
Db	963	UCAAAACAACAGCAUACACCGGCGAGCUCACUGACGCGAGUUCUCCUCCUCCGAGGCGUC	1022
Qy	1022	CCGTGCGAGAAAGTGGGGAATGCTCTCAGTGTGCTGATACCGGTCTCACGAAATGTGGCC	1081
Db	1023	CCAUGUGAGAAUAAUAGGCAACUUGGCGUUGGUGGAGUACAAUAGUACACCCACGUGGCU	1082
Qy	1082	GTGACGCGGCGCGGCGCTTACGAGGGGCTTCCGAGCGGACATGATGATGATGATGATG	1141
Db	1083	GUGAAACACCGGCGGCGCTTACGAGGGGCTTCCGAGCGGACATGATGATGATGATGATG	1142
Qy	1142	TCCGCGAGCTGCTGCTGCGCTTACGTGGGGGACCTCTGCGGTGGGCTGATGCTCGCA	1201
Db	1143	GCAGUACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1202
Qy	1202	GCCCAATGTTCTTCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1261
Db	1203	UCCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1262
Qy	1262	ATCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1321
Db	1263	AUCUACGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1322
Qy	1322	CCCAAGGCTACCATGATGCTTGGGCTAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1381
Db	1323	CCCAAGGCTACCATGATGCTTGGGCTAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1382
Qy	1382	ATTAGCGGCGCTCATTTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1441
Db	1383	AUUAUCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1442
Qy	1442	TGGGCGGAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1501
Db	1443	UGGGCGGAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1502
Qy	1502	GTTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1561
Db	1503	ASCGGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1562
Qy	1562	AGGCGAGAAATCCAGCTGCTGTTAAACCAATGCGAGCTGCGGCGGCGGCGGCGGCGGCGG	1621
Db	1563	AGGCGAGAAATCCAGCTGCTGTTAAACCAATGCGAGCTGCGGCGGCGGCGGCGGCGGCGG	1622
Qy	1622	AACGTCAATGACTCTCTTGCACACCGGCTTTTATCGCGCTCTCTGTTCTTACACCAAGCTTC	1681

7622 TCCTGGTCTACTTGTCTCCGAGGAGCGACTCCGCTGCTGCTCCATGTCTACTCTC 7681
 Db |||||
 7623 TCCTGGTCCACAGTCTCTGATCAAGAGATTCTGTTATCTGCTCTATGTCTACTCTC 7682
 Qy |||||
 7682 TGGACCGGGCTCTAATACTCTCTGTTAGTCCGAAAGAGAGAGTGTACCGATTAACCC 7741
 Db |||||
 7683 TGGACCGGGCTCTAATACTCTCTGTTAGTCCGAAAGAGAGAGTGTACCGATTAACCC 7742
 Qy |||||
 7742 TTGAGCAACTCTCTGTTGCTATATCAACAAGAGTGTACTGTATCAACAAGAGAGCC 7801
 Db |||||
 7743 CTGAGTAATTTCTGCTGATCGGTTCCATAAAGAGTGTACTTCCAAACCTCGAGAGTGC 7802
 Qy |||||
 7802 TCACTAAGGCTTAAAGAGTAACTTTTGTAGTATGCAAGTGTCTCGACTCTCTACTAGC 7861
 Db |||||
 7803 TCTCTGAGGCAAGAGAGTGTACTTTTGAAGAGTGTCTGAGAGTGTCTGAGAGTGTAC 7862
 Qy |||||
 7862 TCAGTCTTAAAGAGCAATTAAGTGTAGCGGCTTCAAGGTCTACCGCAAGGCTCTCAACATG 7921
 Db |||||
 7863 TCAGTCTTGAAGAGCAATTAAGGCTTGAAGCGGCGCTCTAAGGTTGTGAGGCTCTCAAGTA 7922
 Qy |||||
 7922 GAGAGGCTTGCAGTGTAAACCCCAACCCCTTCTGCAAGATCTAAATATGGTTTGGGCT 7981
 Db |||||
 7923 GAGAAAGCTTGCAGGCTGACCCCGCCCACTCCCGCAATTCGCGATACGGAATTTGGGCA 7982
 Qy |||||
 7982 AAGAGCTTCCAGCTTGTCCGGAGGCGGCTTAAACCAATCAAGTCCGTTGCAAGGAC 8041
 Db |||||
 7983 AAGAGGTCGAGCTTATTCAGAGGCGGCTTAAACCAATCCGTTCCGTTGGAGGAC 8042
 Qy |||||
 8042 CTCCTGAGGACTCAGAAACCAATTCACCAACCAATTAAGTCCCAAAATAGAGTGTTC 8101
 Db |||||
 8043 CTCCTGAGGACCAACRTACCCCAATTTGACCAATATCATGCTTAAATAGAGTGTTC 8102
 Qy |||||
 8102 TGGTGGACCCCAAGGAGGCGGCAAGAGAGCTGCGCTTATCTGTTTACCTGACCTC 8161
 Db |||||
 8103 TGCATTTGATCAACTAAGGTGGGAAAGAGCCAGCTCGCTCATCTGTTATACCCGACCTT 8162
 Qy |||||
 8162 GCGTTCAGGCTTCCGAGAGAGTGGCTTTATGACATTACACAAATCTCTCAGGCG 8221
 Db |||||
 8163 GCGTTCAGGCTTCCGAGAGAGTGGCTTTATGACATTCACAAAGCTTCCAAAGCG 8222
 Qy |||||
 8222 GTGATGGGCTTCTTTATGAGTTCAGTATTCCTCCGCTCAGCGGTAGAGTTCTCTTG 8281
 Db |||||
 8223 ATAATGGGCTATCTATGGTTTCCAATCTCTCCGAGAACGGTTCGATTTCTCTCTC 8282
 Qy |||||
 8282 AAGATGGCGGCAAGAGAGCTTATGGGTTTTCGTATGATACCGATGCTTTGAC 8341
 Db |||||
 8283 AAGCTTGGGGAAGTAAAGAGGACCAATGGGTTCTCGTATGACACCCGCTGCTTGAC 8342
 Qy |||||
 8342 TCAACGCTCACTCAGAGAGACATCAGAGCTGAGAGTCCATATATCGGCTGCTCTTG 8401
 Db |||||
 8343 TCAACGCTCAGAGAGGACATAGAACAGAGATCCATATACAGGCTGTTCTCTG 8402
 Qy |||||
 8402 CCGAGAGGCCCACTGCTATACACTCGCTAATCAGAGCTTTACGTGGAGGCGCT 8461
 Db |||||
 8403 CCTCAAGAGCCAGAACTGTCTACACTCGCTCAGTACAGAGCTTTTACGTAGGAGGCC 8462
 Qy |||||
 8462 ATGTTCAAGCAGAGGCCAAACCTGGGGTACAGGCTTCCCGCCAGCGGGTGTCTC 8521
 Db |||||
 8463 ATGACAAACAGCAAGAGGCCAATCTCGGCTACAGGCTTGGCGCGCAAGCGGKGTTC 8522
 Qy |||||
 8522 ACCACTAGCATGGGAACACCATCATGCTAGTGAAGCTTACGGCTTGAAGCT 8581
 Db |||||
 8523 ACCACAGCATGGGAATACCATGATGTTATCATCAAGCCCTTCAGCGTGAAGCT 8582
 Qy |||||
 8582 GCAGGATATCGCGCCCAAACTGCTGATGCGGAGTACTGTTGTTCTCTCAGAA 8641
 Db |||||
 8583 GCRGGATCGTGACCTGTTATGTTGGTGTGGAGAGGACCTGCTGCTCTCAGAG 8642
 Qy |||||
 8642 AGCCAGGACCGAGAGGAGAGCGGACCTCAGAGGCTTACAGAGGCTATGACAGG 8701
 Db |||||
 8643 AGCCAAAGGTAAAGAGGAGGAGCGGCAACCTCAGAGCTTTTACAGGAGGCTATGACAGG 8702

8702 TATTTCTGCCCCCTCTGTTGACCCCGGAGAGTATGATCTGAGCTGTATAACATCT 8761
 Db |||||
 8703 TATTCGCGCCCTCCCGTGAACCTTCCAGAGCGGAATATGACTTGGAGCTTATAACATCC 8762
 Qy |||||
 8762 TGTCTCTCAATATGTTGTGTGGCTGGGCGCCCAAGAGCGGCGGAGATCTACTCTGACC 8821
 Db |||||
 8763 TGTCTCTCAAGCTATCGGTAGCTGAGCTCTCGGGTGGCGCGGCTACTTCTTAACC 8822
 Qy |||||
 8822 AGAGACCTTACCACCTCAATCGCCGGGCTGCTGGGAACAGTTAGACACTCCCTGTC 8881
 Db |||||
 8823 AGAGACCTTACCACCTCAATCGAGCTGCTGGGAACAGTTAGACACTCCCTGTC 8882
 Qy |||||
 8882 AATTCATGCTGGGAAACATCATCAGTACGCGCCGACCATATGGCTCGCATGTCCTG 8941
 Db |||||
 8883 AATTCATGCTGGGAAACATCATCAGTACGCGCCGACCATATGGCTCGCATGTCCTG 8942
 Qy |||||
 8942 ATGACACACTTCTCTCAATCTCATGCTCAAGACACGCTGGAACAGAACTCTCAACTT 9001
 Db |||||
 8943 ATGACTCACTTCTCTCAATCTCATGCTCAAGACACGCTGGAACAGAACTCTCAACTT 9002
 Qy |||||
 9002 GAGATGTACGAGCGGTACTCTCGTGTGCTTGGACCTCCAGCTATAATTTGAAGG 9061
 Db |||||
 9003 GAGATGTACGAGCGGTACTCTCGTGTGCTTGGACCTCCAGCTATAATTTGAAGG 9062
 Qy |||||
 9062 TTACATGGCTTGAAGCTTTTCTCTGACACATACACTCCCGACGAACTGACAGCGGTG 9121
 Db |||||
 9063 CTACATGGCTTGAAGCTTTTCTCTGACACATACACTCCCGACGAACTCTCAGCGGTG 9122
 Qy |||||
 9122 GCTTCAGCGCTCAGAAACTTGGGCGGCGCCCTCAGAGCTGGAAGAGCGGCGACGT 9181
 Db |||||
 9123 GCAGCAACTCTCAGAAACTTGGGCGGCGCTCCCTTAGAGCGTGGAAAGTGGGCGGT 9182
 Qy |||||
 9182 CGAGTCAAGCGCTCTCATCTCCGTTGGGCGGAGCGCGCTTGGCTGATATCTC 9241
 Db |||||
 9183 GCGGTGAGAGCTTCACTCATCGCCCAAGAGCGGCGGCTTGTGGCGCTACCTC 9242
 Qy |||||
 9242 TTCAATTTGGGCGGTGAAGACCAAGCTCAAACTCACTTCCCGGAGCGGCGCTCTCTG 9301
 Db |||||
 9243 TTCAATTTGGGCGGTGAAGACCAAGCTCAAACTCACTTCCCGGAGCGGCGCTCTCTG 9302
 Qy |||||
 9302 GATTATTCAGTGTGTTACCTCGGCGGCGGCGGCGGCGCAATTTATCAGAGGTGTG 9361
 Db |||||
 9303 GATTATTCAGTGTGTTACCTCGGCGGCGGCGGCGGCGCAATTTATCAGAGGTGTG 9362
 Qy |||||
 9362 CCGTCCGACCGGCTTATGCTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 9421
 Db |||||
 9363 CAGTCGACCGGCTTATCT 9422
 Qy |||||
 9422 TTCTCTACTCCCGCTCGGTAGAGCGGCAACATTTAGCTACACTTCCATAGCTAACTGCC 9481
 Db |||||
 9423 TTTTCTACTCCCGCTCGGTAGAGCGGCAACATTTAGCTACACTTCCATAGCTAACTGCC 9482
 Qy |||||
 9482 TTTTCTACTCCCGCTCGGTAGAGCGGCAACATTTAGCTACACTTCCATAGCTAACTGCC 9510
 Db |||||
 9483 TTTTCTACTCCCGCTCGGTAGAGCGGCAACATTTAGCTACACTTCCATAGCTAACTGCC 9511

RESULT 4

US-07-925-695-6

; Sequence 6, Application US/07925695

; Patent No. 5428145

; GENERAL INFORMATION:

; APPLICANT: OKAMOTO, Hiroaki

; APPLICANT: NAKAMURA, Tetsuo

; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOME,

; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, ANTIGEN, ANTIBODY AND

; TITLE OF INVENTION: DETECTION SYSTEMS

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSES: Beveridge, DeGrandi, Weilacher & Young

; STREET: 1850 M Street, N.W., Suite 800

; CITY: Washington

; STATE: D.C.

6483	ATGAAATAACAGCCCGAAGACTTGTCTGAACTCTGTGGCAGGGGACTTTTCCCCCATTAAT	5543
6542	TGTTACACGAGGGCCAGTGCCTGCCAAACCCCGCCAAACTTTAAAGTTCGCCAATCTGG	6601
6543	TGTTACACAGAAAGGCGVTCGTGCGCAAAACCCCTCTTAATTACAAGACCGCAATTTGG	6602
6602	AGGTTGGCGGCTCAGAGTAGCGGAGGTGACGAGCAGCAGCGGTATACCACTACATAACA	6661
6603	AGGTTGGCAGCGTCGGAGTAGTTGAGGTACACAGCATGGCTCTTTCTCGTAGTGAACR	6662
6662	GGACTCACCACTGATACTTTGAAAGTCCCTTCGCCAACTACCTCTCTCCCGAGTCTTTTCC	6721
6663	GGGTTAACCACTGACAACTTAAGTTCCTTGCACAGGTACCACTCCAGAAATTTTCTCT	6722
6722	TGGGTGACGAGTAGCAGATCCATAGTTTGGCCCCCACACCGAAGCGGTTTTTTCGGGAT	6781
6723	TGGGTGACGCGGTGCAATCCACCGATTGGCCCCCGTWCACAGGTCCCTTTCTTTCGGGAT	6782
6782	GAGTCTCTGTTCTGCTTGGCTTAATTCATTTGTCGTCGGGTCCCAAGCTTCCCTTGCAC	6841
6783	GAGGTAAAGTTTCAACCGTAGGCTTAATCTCTCTGTCGTCGGCTCTCAGCTCCCTTGCAT	6842
6842	CCTGAACCCGACACAGACGTATTGATGTCTCATGTCTAAACAGATCCAATCTCATACACGGCG	6901
6843	CCCTGAGCGGACACCGAGTACTGGCTCTATGTTGACAGACCCGTCCACATCACACGGK	6902
6902	GAGACTCAGCGCGCGCTTTAGCGCGGGGTCAACCCCATCCGAGGCAAGCTCTCTCGGCG	6961
6903	GAGCGCGACGCGCGATTGGCAAGGGATCTCCCCCYTCACAGGCTAGCTCTCTCAGCG	6962
6962	AGCCAGCTATCGGCACACATCGCTCGGAGCCTCTGACACACCGCAAGGCTATGAT	7021
6963	AGCCAGCTCTCTGCCCCGCTCTTGAAGGCTACCTGTGTACACCCCAATGAACACGACATGAT	7022
7022	GTGACATGTTGGATGCTAACTCTGTCATGGGGCGCATGTGACTCGCATAGAGTCTGG	7081
7023	TGTGACATGTTGGATGCTAACTCTTTCATGGAGGAGTGTGATCCGAGTTGAGTCTGAC	7082
7082	TCCAAGTGGTCTGGAATCTCTCGACCCAATGGTCTGAAGAAAGGACGCACTTGAG	7141
7083	TCTAAGGTGATCGTTCTAGACTCTCCCTGATTCCATGACTGAGGTAGAGGATGATCGTAG	7142
7142	CTTTCGATACCATCAGAAATACATGCTCCCAAGAAAGAGTTTCCACACAGCTTTTACCGGCC	7201
7143	CTTCTGTATCACTCAGAGTACTCTGATCAAGAGGAGAAATTTCCCAACCGCGCTGCTCTCT	7202
7202	TGGCAGCGCTGATTACAAACCCAGCTGTGGAATCGTGAAAGGCGCAGATTACCAA	7261
7203	TGGGCGCTGCAGACTACATCTGTTTGTGAGACATGGAAGAGCGCGGCTATGAA	7262
7262	CCGCGCACTGTGGCGGTGTGCTCTCCCTCTCTCTAGGAAACCCCGACGCTCTCCCCA	7321
7263	CCACCACTGTCTAGGCTGTGCCCTCCCCCCACACACACTCAAAACGCGAGTCTTCACT	7322
7322	AGGAGGCGCCGACAGTGGGCTTAAGTAGAGACTCCATAGGAGATGCCCCCTTCAAAGCTG	7381
7323	CGGAGGCGCGCGCYAARTCTCTGACGAGACRATGTGTGAAGGGRTCTCTCAGGGAGATG	7382
7382	GCCATTAAAGTCTTTGGCCAGCCCCCCCCAAGCGCGATTTCAGGGCTTTCACCGGGCG	7441
7383	GCTGACAAAGTRCTCAGCGCTCTCCAGACAAATGATCTCGGTCTACTCTCAGCTGGAGCG	7442
7442	GGCGCTGCCGATTCCGGCAGTCAGACGCTCTCTGATGAGTTGGCCCTTTCGGAGACAGGT	7501
7443	GATACCGGAGAGACATCGTCCAGCAACCTCTGACGAGACTGCGCTTCAGAACGGGG	7502
7502	TCCATCTCTTCCATGCCCCCTTCGAGGGGAGCTTTGGAGATCCAGACTCGAGCCTGAG	7561
7503	TCAGTCTCTCATGCTCTCCCTTCGAGGGAGAGCCGGGAGACCCYGAACCTCGAGTTTGA	7562
7562	CAGGTAGAGCCCCCAACCCCCCCCCCAGGGGGGTGGCAGCTCCCGGCTCGCACTCGGG	7621
7563	CCAGTGGGATCCGCTTCCCTTTCGAGGGGAGTGTGAGGTCAATTGATTCGCACTAAG	7622

Db 1023 CCATGTCAGAAATGATTAATGACACCTTGTGCTTGTGATACAAAGTAAACACCCACGCTGGCT 1082
 Qy 1082 GTGAGCGGCGCGCGCGCTTACCGAGGCTTGGGAGCGACATCGACATGCTTGTGATG 1141
 Db 1083 GTGAAACACCGCGGCTGCTCACTCGTAGCTCGGAAACACACGTCGACATGATCGTAATG 1142
 Qy 1142 TCGGCCACGCTCTGCTCGCCCTTACGCTGGGGACCTCTGCGGTGGGGTGTGCTCGCA 1201
 Db 1143 GCAGCTACGCGCTGCTGCGCTTGTATGTGGAGATGTGTGGGGCGGTGATGATYCTA 1202
 Qy 1202 GCCCAATGTCATGTCCTCGCGCAGCACACCTGTTTGTCCAAGACTGCAATGTCTCC 1261
 Db 1203 TCGCAGGCTTTCATGGTATCACCAACGCGCAACACTTCCACCAAGAGTGCACTGTTC 1262
 Qy 1262 ATCTACCTGGTACCATCTGACGACCGATGGCATGGGACATGATGATGAACTGGTCG 1321
 Db 1263 ATCTACCAAGGTACATCACCGGCTATCGATGGCATGGGACATGATGCTFRACCTGTCT 1322
 Qy 1322 CCCACGCTTACCATGATCTTGGCGTACCGATGCTGCTCCCGAGGTCAATTATAGATC 1381
 Db 1323 CCAACTCTTTCATGATCTCGCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1382
 Qy 1382 ATTAGCGGGCTCATTTGGGGCTCATGTTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1441
 Db 1383 ATVTTCGGCGGCAATGGGGTGTGGYTTVGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGG 1442
 Qy 1442 TGGGGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1501
 Db 1443 TGGGCCAAAGTCRTYGGCATCTCTCTTCTTGTGGGGAGTGGATGCAWCCACCTATTC 1502
 Qy 1502 GTTGGGGTCTTGGCGGCGAGACACCGGCGCTCACAGCTTATTTGACATGGGCGCC 1561
 Db 1503 ASCGGYAGSNAGGGGTGTRCCGCKMGGGTCTKTRGGCTTATTAMTACTGGTGCC 1562
 Qy 1562 AGGCAGAAATCCAGCTGTTTAAACCAATGGAGCTGGGACATCAACCGCACCGCCCTG 1621
 Db 1563 AAGCAGAACTCTATTTTATCAACCAATGGAGCTGGGACATCAACCGCACCGCTGCTC 1622
 Qy 1622 AACTGCAATGCTCTTGGCACACCGGCTTATCGGCTCTCTGCTTCTACACCCACGCTTC 1681
 Db 1623 AATTGCAATGACAGCTASAGAGGGGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1682
 Qy 1682 AACTCGTCAGGATGCTCCGAAAGCATGTCGCTGCGCAGTATCGAGGCTTCCGGGTG 1741
 Db 1683 AACAGCTCTGGCTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1742
 Qy 1742 GGATGGGCGCTTGGCAATATGAGGATAATGTCACCAATCCAGAGGATATGAGACCTAT 1801
 Db 1743 GGCTGGGGAACCTTTGGAATACGAAACCAACGTCACCAAGATGAGGACATGAGGCGGTAC 1802
 Qy 1802 TGCTGGCACTACCCACCAAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1861
 Db 1803 TGCTGGCAATTAACCCCGGAGGCTTGGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1862
 Qy 1862 GTGTACTGTTTACCCCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1921
 Db 1863 GTCTATTGTTTACCCCTAGCCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1922
 Qy 1922 ACTTACAGCTGGGGGAGATGAGACAGATGCTTCTTCTTATTTGAAACAGACCTCGACACCG 1981
 Db 1923 ACCTACACCTGGGGGAGAAACGAGACCGATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1982
 Qy 1982 CTGGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2041
 Db 1983 CGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2042
 Qy 2042 GCACACCTCTCGCTACTAGAGCTGACTTCAACCGCAGCAGGACCTGCTGCTGCTGCTGCTGCT 2101
 Db 2043 GCACACCTCTCGCTACTAGAGCTGACTTCAACCGCAGCAGGACCTGCTGCTGCTGCTGCTGCT 2102
 Qy 2102 GACTGTTTTAGGAAGCATCTGATACCACTTACCTCAAAATCGGCTCTGGGCGCTTGGCTC 2161
 Db 2161

Db 2103 GACTGTTTTAGGAAGCACCCAGATGCTACTTATCTTAAAGTGGAGCAGGCGCTTGGTTA 2162
 Qy 2162 AGCCCAAGTGTGCTGATCGACTACCCCTAGGCTCTGGCATTAACCCCTGACAGTTAAC 2221
 Db 2163 ACTCCAGTGTGCTGATAGCTACCCCTTATAGRYTGTGSCATTAATCCGTCGACTGTAAAC 2222
 Qy 2222 TATACCATCTTCAAAATAAGGATGATGTGGAGGGTGTGAGCAGCAGGCTCACGGCTGCA 2281
 Db 2223 TTCAACCATCTTAAGCGCGGATGATGTAGGAGGGTGTGAGCAGCAGGCTTCTCCGAGCA 2282
 Qy 2282 TGCAATTTCACTCGTGGGATCGTTGCAACTTGGAGGACAGAGACAGAGTCACTGTCT 2341
 Db 2283 TGCNACTTCAACCGCGGATCGCTGACACTGGAAGATAGGATAGGAGTCACTGTCT 2342
 Qy 2342 CTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2401
 Db 2343 CCACTGCTGATCTCACTACTGAGTGGGGGTGCTCCATGCTCTCTCTGACTTACCA 2402
 Qy 2402 GCCTTGTGCACTGCTTCTTCCACCTCCACCAAAACATCGTGGAGCTTCAATTCATGTAT 2461
 Db 2403 GCACTATCCACTGGCTTATGCACTTCCACCAAAACATCGTGGAGCTGCACTACCTTATC 2462
 Qy 2462 GGCCTATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2521
 Db 2463 GGACTTCTCCGCTCTGACAGATACATCGTGAAGTGGAGTGGGATGCTCTCTCTTTC 2522
 Qy 2522 CTGCTCTTACCGGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2581
 Db 2523 TTGTTGTGGCAGCGCAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2582
 Qy 2582 GCCGAAGCAGCAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2641
 Db 2583 GCCGAAGCAGCGCTTGAAGCTCATCATCTTGGACTCGCTAGCTGCTGCTGCTGCTGCTGCT 2642
 Qy 2642 GGCCTCTATATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2701
 Db 2643 GGTCCGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2702
 Qy 2702 CCTTACGCTTATTTCCCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2761
 Db 2703 CCGTGGCAGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2762
 Qy 2762 CCCCAACAGCTTATGCTTATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2821
 Db 2763 CCACAGCAGGCTTATGCTTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2822
 Qy 2822 GTAATGATCACTCTTCTTACTCTCAACCGGCTTAAAGACCTTCTCAAGCGGTTTTG 2881
 Db 2823 GTAATTATATCCATCTTCTTACTCTTACCCAGCATCAAGATCTCTCTGAGCGGCTTCACTG 2882
 Qy 2882 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2941
 Db 2883 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2942
 Qy 2942 ATCAGGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3001
 Db 2943 CTGAGGTCCGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3002
 Qy 3002 GTGGTGTGACATAACCAAGTGGCTTGGCGGCTTGGCGGCTGCTTACCTCTCTTAA 3061
 Db 3003 CTTGTTGTTGAGGTACGAAATGTTGTTAGCAATCTCTGGGGCTGCTTACCTCTCTTARA 3062
 Qy 3062 GGTGCTTGGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3121
 Db 3063 GCGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3122
 Qy 3122 ATGCAAGGCTCTGCGGGGCGAGTACGCTCCAGATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3181
 Db 3123 CTGGTGAACACCTCGCGGGCTGAGTATCCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3182
 Qy 3182 TGCACTGGCACTTACATCTATGACCACTTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3241
 Db 3183 TGGACCGGCACTTACATCTACGACCACT 3242

Db 9241 CTTCAACTGGCGGTGAAGCAAGCTCAAACTCACTCCATTGCGGAAGCGCGCTCCT 9300
Qy 9301 GGATTTATCAGCTGGTTACCGTGGCGCGCGGGGCGACATTTATCACAGCGTGC 9360
Db 9301 GGATTTATCAGCTGGTTACCGTGGCGCGCGGGGCGACATTTATCACAGCGTGC 9360
Qy 9361 GCGTCCCGGACCCCGCTTATTGCTCTTTGGCCCTACTCTCTTTTGTAGGGGTAGGCCCT 9420
Db 9361 GCGTCCCGGACCCCGCTTATTGCTCTTTGGCCCTACTCTCTTTTGTAGGGGTAGGCCCT 9420
Qy 9421 TTTCTACTCCCGCTCGGTAGAGCGGCAACATTAGCTACATCCATAGCTAACTGTGC 9480
Db 9421 TTTCTACTCCCGCTCGGTAGAGCGGCAACATTAGCTACATCCATAGCTAACTGTGC 9480
Qy 9481 CTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9540
Db 9481 CTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9540
Qy 9541 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9588
Db 9541 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9588

RESULT 3
US-07-925-695-7
Sequence 7, Application US/07925695
Patent No. 5428145
GENERAL INFORMATION:
APPLICANT: OKAMOTO, Hiroaki
APPLICANT: NAKAMURA, Tetsuo
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOME,
POLYNUCLEOTIDES, POLYPEPTIDES, ANTIGEN, ANTIBODY AND
DETECTION SYSTEMS
TITLE OF INVENTION: DETECTION SYSTEMS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Beveridge, DeGrandi, Wellacher & Young
STREET: 1850 M Street, N.W., Suite 800
CITY: Washington
STATE: D.C.
COUNTRY: US
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/925,695
FILING DATE: 19920807
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 287402/91
FILING DATE: 09-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 360441/91
FILING DATE: 05-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Wellacher, Robert G.
REGISTRATION NUMBER: 20,531
REFERENCE/DOCKET NUMBER: 06/87-48009
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2811
TELEFAX: (202) 659-1462
TELEX: WUI 64470
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 9511 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
US-07-925-695-7

Query Match 62.1%; Score 6031.4; DB 1; Length 9511;
Best Local Similarity 76.3%; Pred. No. 0;
Matches 7257; Conservative 105; Mismatches 2147; Indels 0; Gaps 0;
Qy 2 CCGCGCCCTAATAGGGGCGACACTCGCCATGAATCACTCCCTGTGAGGAATCTACTGTC 61
Db 3 CCGCGCCCTGATGGGGGCGACACTCGCCATGAATCACTCCCTGTGAGGAATCTACTGTC 62
Qy 62 TTTACGAGAAACGCTTAGCCATGGCTTAGTATGAGTGTCTAGAGCTTACAGCTCCAGGCC 121
Db 63 TTTACGAGAAACGCTTAGCCATGGCTTAGTATGAGTGTCTAGAGCTTACAGCTCCAGGCC 122
Qy 122 CCGCTCCCGGAGAGCCTAGTGGTCTGCGAAACCGGTGAGTACACCGAAATTCGCGGA 181
Db 123 CCGCTCCCGGAGAGCCTAGTGGTCTGCGAAACCGGTGAGTACACCGAAATTCGCGGA 182
Qy 182 AGACTGGGTCTTTCTTGGATAAACCCACTCTATGCTCCGCTCAVTTGGGCGTGCCTGAT 241
Db 183 AGACTGGGTCTTTCTTGGATAAACCCACTCTATGCTCCGCTCAVTTGGGCGTGCCTGAT 242
Qy 242 AGACTGGGTCTTTCTTGGATAAACCCACTCTATGCTCCGCTCAVTTGGGCGTGCCTGAT 301
Db 243 AGACTGGGTCTTTCTTGGATAAACCCACTCTATGCTCCGCTCAVTTGGGCGTGCCTGAT 302
Qy 302 GCTTGGAGTGCCTCGGGAGGTCTCGTAGACCGTGCACCATGAGCACAAATCCTAAACCT 361
Db 303 RCTTGGAGTGCCTCGGGAGGTCTCGTAGACCGTGCATCATGAGCACAAATCCTAAACCT 362
Qy 362 CAAAGAAACCCAAAGAAACACCAACCGTTCGCCCAACAGAGCTTTAAGTTTCGGGCGGC 421
Db 363 CAAAGAAACCCAAAGAAACACCAACCGTTCGCCCAACAGAGCTTTAAGTTTCGGGCGGC 422
Qy 422 GCGCAGATCGTTGGCGGAGTACTTGTTCGCCCGCAGGGGCCAGAGTTGGGTGGCGC 481
Db 423 GGTCAATCGTTGGCGGAGTACTTGTTCGCCCGCAGGGGCCAGAGTTGGGTGGCGC 482
Qy 482 GCGCAAGGAAGACTTTCGGAGCGGTCCAGCAACGCTGGAAGCGCGCAGCCATCCCTAAA 541
Db 483 GCGCAAGGAAGACTTTCGAGCGATCCAGCGCGTGGAGCGCGCAGCCATCCCTAAA 542
Qy 542 GATCGCGCTCCACTGGGAAATCTTGGGAAACCAAGGATACCTTGGCGCTTATACGGG 601
Db 543 GATCGCGCTCCACTGGGAAATCTTGGGAAACCAAGGATATCTTGGCGCTTATACGGG 602
Qy 602 AATGAGGACTCGGCTGGGAGAGTCTCTGTCGCCCGCAGAGTTCCGCTCCCTCTGG 661
Db 603 AACGAGGTTGCGGCTGGGCGGTGGCTCTCTGTCGCCCGCAGAGTTCCGCTCCCTCTGG 662
Qy 662 GCGCCCAATGACCCCGCATAGGTTCGGGCAACGCTGGGTAAAGTCAATCATACCTTAA 721
Db 663 GCGCCCAATGACCCCGCATAGGTTCGGGCAACGCTGGGTAAAGTCAATCATACCTTAA 722
Qy 722 TCGCGCTTTCGCACTCATGGGTACATCTCTGTCGGGCGCGCGCTCCGCGGCGTGC 781
Db 723 TGTGGTTTTCGCACTCATGGGTACATCTCTGTCGGGCGCGCGCTCCGCGGCGTGC 782
Qy 782 GCGAGAGTCTCGCGCATGGCGTGGAGTCTCTGGAGGCGGGTTAATTTTTCACAGG 841
Db 783 GCGAGAGTCTCGCGCATGGCGTGGAGTCTCTGGAGGCGGGTTAATTTTTCACAGG 842
Qy 842 AACTTTACCGGTTGCTCTCTTTTCTATCTTCTGTCGGCGCTCTCTCTGCAATCAC 901
Db 843 AATTTACCGGTTGCTCTCTTTTCTATCTTCTGTCGGCGCTCTCTCTGCAATCAC 902
Qy 902 CCGGTCTCGCTGCGGAAGTGAAGAACATCAGTACCGGCTACATGGTGTACATCAAG 961
Db 903 CCAGTGTCTGCACTGGAAAGTCAAGAACATYAGTTCTAGCTACTACGCCCACTAAT 962
Qy 962 ACCAATGACAGCATTTACCTGGCAGCTCCAGGCTGTCTCTCCAGTCCCGGGTGC 1021
Db 963 TCAAAACACAGCATCACCTGGGAGTCTACTGACGAGTTCTCTCTCTCTGATGCG 1022
Qy 1022 CCGTGGGAAAGTGGGAAATGCATCTCAGTGTGGATACCGGCTCTCACGAATGTG 1081

4861	DB		CACTGGTGAGCGAGCCTCAGGAATGTTTGACAGTGTAGTGCTCTGCGAGTGTACGATGC	4920	DB		CGTCGCAATTCAGAGTCAATGCTCTGGCGAGAAGCCCTCCATGAGGAGTGTGTCTCAACCTGCT	6000
4921	QY		AGGGGCGCATGATGAGTGTACACCATCGGAGACACCGTCAAGGTCAAGGCGTATTT	4980	QY		GCCTGGAATTCCTGCTCCGGGTGCGCTGTGTGGAGTCACTGCGCGCCATCTCTGG	6060
4921	DB		AGGGGCGCATGATGAGTGTACACCATCGGAGACACCGTCAAGGTCAAGGCGTATTT	4980	DB		GCCTGGAATTCCTGCTCCGGGTGCGCTGTGTGGAGTCACTGCGCGCCATCTCTGG	6060
4981	QY		CAACAGCCCGTGTGCTGTGCGCAAGACCATCTTGAGTTTGGGAGGACGTTTTCAC	5040	QY		CCGACACGTCGGGAGGCGCGCTCCAAATGGAATAGACTCAATGCTCTGCTTTGC	6120
4981	DB		CAACAGCCCGTGTGCTGTGCGCAAGACCATCTTGAGTTTGGGAGGACGTTTTCAC	5040	DB		CCGACACGTCGGGAGGCGCGCTCCAAATGGAATAGACTCAATGCTCTGCTTTGC	6120
5041	QY		CGGCTCACACATAGATCCCATCTTCTTCCCAAAACAAAGCAATCGGGGAAATTT	5100	QY		TTCCAGAGAAATCAAGTCGCGCCCAACCACTACGTCGAGTCGATCGCTCGCAGCG	6180
5041	DB		CGGCTCACACATAGATCCCATCTTCTTCCCAAAACAAAGCAATCGGGGAAATTT	5100	DB		TTCCAGAGAAATCAAGTCGCGCCCAACCACTACGTCGAGTCGATCGCTCGCAGCG	6180
5101	QY		CGCATCTTAACAGCCTACAGGCTACAGTGTGGGTAGGGCCAAAGCCCGCCCGCTC	5160	QY		TGTCGCCCAACTACTTGGCTCCCTTACCATACCAAGCTCTCAGNAGACTCCACAATG	6240
5101	DB		CGCATCTTAACAGCCTACAGGCTACAGTGTGGGTAGGGCCAAAGCCCGCCCGCTC	5160	DB		TGTCGCCCAACTACTTGGCTCCCTTACCATACCAAGCTCTCAGNAGACTCCACAATG	6240
5161	QY		CTGGAGCTCATGTGGAAGTGTGACTCGACTCAAGCCACACTCGTGGGCGCCACACC	5220	QY		GATTACTGAGGACTGCGCCCATCCCATCGCGCTCGTGGCTCCGCGATGTGTGGACTG	6300
5161	DB		CTGGAGCTCATGTGGAAGTGTGACTCGACTCAAGCCACACTCGTGGGCGCCACACC	5220	DB		GATTACTGAGGACTGCGCCCATCCCATCGCGCTCGTGGCTCCGCGATGTGTGGACTG	6300
5221	QY		TCTCTGTACCGCTTGGGCTCTGTTACCAACGAGGTCAACCCTCACATCCCGTGACAA	5280	QY		GGTTTGACCATCTTAAAGATTTTAAATTTGGCTGACCTCCAAATTAATTCCTCAAGAT	6360
5221	DB		TCTCTGTACCGCTTGGGCTCTGTTACCAACGAGGTCAACCCTCACATCCCGTGACAA	5280	DB		GGTTTGACCATCTTAAAGATTTTAAATTTGGCTGACCTCCAAATTAATTCCTCAAGAT	6360
5281	QY		ATACATCGCACCTGATGACGCGGACTGTGAGTGTACGACGACATCGGTCTTGCC	5340	QY		GCCCGGCTCCCTTGTCTCTGTCGCAAGGGGTACAGGGGTGTGGCGGCACTGG	6420
5281	DB		ATACATCGCACCTGATGACGCGGACTGTGAGTGTACGACGACATCGGTCTTGCC	5340	DB		GCCCGGCTCCCTTGTCTCTGTCGCAAGGGGTACAGGGGTGTGGCGGCACTGG	6420
5341	QY		AGGGGAGTCTTGCGGCGCTCGCGGTATTCGCTGCGGACCGGCTGTGTGATCAT	5400	QY		CATCATGACACACGCGTGTCTGGCGCCCAATATCTCTGCAATGTCCGCTTGGGCTC	6480
5341	DB		AGGGGAGTCTTGCGGCGCTCGCGGTATTCGCTGCGGACCGGCTGTGTGATCAT	5400	DB		CATCATGACACACGCGTGTCTGGCGCCCAATATCTCTGCAATGTCCGCTTGGGCTC	6480
5401	QY		CGGCGCTTGACATTAACAGGAGCGCTGTTGCGCGGACCAAGAGGTCTCTATGA	5460	QY		CATGAGAATCACGGGCTTAAGACCTGATGAATATCTGGCAGGGGACCTTTCCATCAA	6540
5401	DB		CGGCGCTTGACATTAACAGGAGCGCTGTTGCGCGGACCAAGAGGTCTCTATGA	5460	DB		CATGAGAATTAACGGGCTTAAGACCTGATGAATATCTGGCAGGGGACCTTTCCATCAA	6540
5461	QY		GGCTTTGATGAGTGAAGTGTGCTCTAGGGGCTCTCATTAAGAGGGGCGAGG	5520	QY		TTGTTACAGGAGGCGAGTGTGCGGAAACCGCGGCCAAACTTTAAGGTGCGCATCTG	6600
5461	DB		GGCTTTGATGAGTGAAGTGTGCTCTAGGGGCTCTCATTAAGAGGGGCGAGG	5520	DB		TTGTTACAGGAGGCGAGTGTGCGGAAACCGCGGCCAAACTTTAAGGTGCGCATCTG	6600
5521	QY		GATAGCGAGATCTGAAGTCCAAGATCCAAAGCTTATTCAGCAAGCTTCCAAACAGC	5580	QY		GAGGTGCGGCTCAGAGTACGCGAGGTGACGAGCACGGGTATACCACTTACATAAC	6660
5521	DB		GATAGCGAGATCTGAAGTCCAAGATCCAAAGCTTATTCAGCAAGCTTCCAAACAGC	5580	DB		GAGGTGCGGCTCAGAGTACGCGAGGTGACGAGCACGGGTATACCACTTACATAAC	6660
5581	QY		TCAAGACATACACCCGCTGTGAGGCTTCATGCGCCCAAGGTAGAACATTTCTGGGCCAA	5640	QY		AGGACTACCACTGATTAATTTGAAAGTCCCTGCGCAACTACCTCTCCGAGTCTTTTC	6720
5581	DB		TCAAGACATACACCCGCTGTGAGGCTTCATGCGCCCAAGGTAGAACATTTCTGGGCCAA	5640	DB		AGGACTTACCACTGATTAATTTGAAAGTCCCTGCGCAACTACCTCTCCGAGTCTTTTC	6720
5641	QY		ACATGAGGAACTTCAATACGCGCATTAATACCTCGCAGGACTTCAACACTGCGCAG	5700	QY		CTGGGTGAGGAGTGCAATCCATAGTTTGGCGCCCAACCGAGCGGTTTTCGGGA	6780
5641	DB		ACATGAGGAACTTCAATACGCGCATTAATACCTCGCAGGACTTCAACACTGCGCAG	5700	DB		CTGGGTGAGGAGTGCAATCCATAGTTTGGCGCCCAACCGAGCGGTTTTCGGGA	6780
5701	QY		GAACCTGAGTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT	5760	QY		TGAGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	6840
5701	DB		GAACCTGAGTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT	5760	DB		TGAGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	6840
5761	QY		AAGCACCACATCTCTTCAACATTTTGGGGGCTGCTAGCATCCCAATTTGCACACC	5820	QY		CCCTGAACCCGACACAGAGTATGATGTCATGTCGTCGTCGTCGTCGTCGTCGTCGTCG	6900
5761	DB		AAGCACCACATCTCTTCAACATTTTGGGGGCTGCTAGCATCCCAATTTGCACACC	5820	DB		CCCTGAACCCGACACAGAGTATGATGTCATGTCGTCGTCGTCGTCGTCGTCGTCGTCG	6900
5821	QY		CGGGGGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	5880	QY		CGGACTGAGGAGTGCAATCCATAGTTTGGCGCCCAACCGAGCGGCTCTCCGCG	6960
5821	DB		CGGGGGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	5880	DB		CGGACTGAGGAGTGCAATCCATAGTTTGGCGCCCAACCGAGCGGCTCTCCGCG	6960
5881	QY		CTTAGGTAGGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT	5940	QY		GAGCAGCTTATCGGACCATCGCTGCGAGCCACTGCAACCCGAGCGGAGCTCTATGA	7020
5881	DB		CTTAGGTAGGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT	5940	DB		GAGCAGCTTATCGGACCATCGCTGCGAGCCACTGCAACCCGAGCGGAGCTCTATGA	7020
5941	QY		CGTCGCAATTCAGATCATGCTGCGCGAGAGCCCTCCATGAGGAGTGTGCTCAACTGCT	6000	QY		TGTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	7080
5941	DB		CGTCGCAATTCAGATCATGCTGCGCGAGAGCCCTCCATGAGGAGTGTGCTCAACTGCT	6000	DB		TGTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	7080

Qy	2701	CCCCCTAGCTACCTATTCCCTCACTGGCCCTGCTCTTTAGCTACTGCTCTAGCACTT	2760	3781	AAGACGCGGGGAGAAACGGGAGCGCTACTCTCCCGAGACCTCTTTCCACCTTGAAGG	3840
Db	2701	CCCCCTGGCTACTTATTTCCCTCACTGGCCCTATGCTCTTTGGCTACTGCTCTAGCACTT	2760	3781	AAGACGCGGGGAGAAACGGGAGCGCTACTCTCTCCCGAGACCTCTTTCCACCTTGAAGG	3840
Qy	2761	GCCCCAACAGGCTTATGCTTTATGACGATCTGTGCTATGCGCAGATAGGAGCGCTCTGCT	2820	3841	GTCTCAGAGGAGCGCGCTGCTATGCCCGAGGGGCGACGCTGTGCGAGTCTTCCGGGCGAG	3900
Db	2761	GCCCCAACAGGCTTATGCTTTATGAGGCTATGCTATGCTGCTGCTGCTGCTGCTGCT	2820	3841	GTCTCAGAGGAGCGCGCTGCTATGCCCGAGGGGCGACGCTGTGCGGGGTCTTCCGGGCGAG	3900
Qy	2821	GGTAATGATCACTCTCTTTACTCTCACCCCGGGTATAGACCTTCTAGCGGGTTTTT	2880	3901	TGTGTGCTCTCGGGGCGTGGCTAAGTCCATAGATTTCCTCCGCTTGAGACACTCCGACAT	3960
Db	2821	GGTACTGATCACTCTCTTTACTCTCACCCCGGGTATAGACCTTCTAGCGGGTTTTT	2880	3901	CGTGTGCTCTCGGGGCGTGGCTAAGTCCATAGATTTCCTCCGCTTGAGACACTCCGACAT	3960
Qy	2881	GTGCTGCTGCTGCTATCTCTGACCTCTGGGGAGCTATGGTCCAGAGTGGGCAACACC	2940	3961	CGTCAACGCGCTCCCCCACCCTTTAGTGAACAAAGACACACCTGCTGTGCCAGACTTA	4020
Db	2881	GTGCTGCTGCTGCTATCTCTGACCTCTGGGGAGCTATGGTCCAGAGTGGGCAACACC	2940	3961	CGTCACTCGGCTCCCCCACCCTTTAGTGAACAAAGACACACCTGCTGTGCCAGACTTA	4020
Qy	2941	TATGAGGCTGGGGTGGCGCTGATGCGATCATATGGGCGCTGCCCATATTCCTACCCAGG	3000	4021	TCAGTTCGGGTACTTGCATGCCCGGCTGGCAGTGGAGAGACCAAGATTCTCTGCTCGC	4080
Db	2941	TATGAGGCTGGCGGCTGGCGCTGATGCGATCATATGGGCGCTGCCCATATTCCTACCCAGG	3000	4021	TCAGTTCGGGTACTTGCATGCCCGGCTGGCAGTGGAGAGACCAAGATTCTCTGCTCGC	4080
Qy	3001	TGTGGTGTGACATACCAAGTGGCTTTGGCGGCTGCTTGGGCTGCTTACCTCTTAAA	3060	4081	ATATGCTGCTCAGGGGTATAAAGTGTAGTGTCTTAATCCCTCAGTGGCTGCCACCTCTGGG	4140
Db	3001	TGTGGTGTGACATACCAAGTGGCTTTGGCGGCTGCTTGGGCTGCTTATCTCTTAAA	3060	4081	GTATGCGGCTCAGGGGTACAAAGTGTAGTGTCTTAATCCCTCAGTGGCTGCCACCTCTGGG	4140
Qy	3061	AGGTGCTTTGACGCGCGCTGCTACTTCTGTCAGGGCTCAGCTCTACTGAGGATGTGCAC	3120	4141	GTTTGGGGGCTACTTGTCTTAAGGACATGCGCATCAATCCAACTTAGGACTGGGGTCTAG	4200
Db	3061	AGGTGCTTTGACGCGCGCTGCTACTTCTGTCAGGGCTCAGCTCTACTAAGGATGTGCAC	3120	4141	GTTTGGGGGCTACTTGTCTTAAGGACATGCGCATCAATCCAACTTAGGACTGGGGTCTAG	4200
Qy	3121	CATGGCAAGGCTATCGCGGGGGGAGTGTGCTCAGATGGCGCTACTAGCCCTTGGGAG	3180	4201	GACTGTGACGACCGGGGCGCCCTACGCTACTCTCCATATGSCAAATTCCTCGCCCATGG	4260
Db	3121	CATGGTAAGGCTATCGCGGGGGTGTAGTGTGCTCAGATGGCGCTACTAGCCCTTGGGAG	3180	4201	GACTGTGACGACCGGGGCGCCCTACGCTACTCTCCATATGSCAAATTCCTCGCCCATGG	4260
Qy	3181	GTGAGCTGGCACTTATCATATGACACCTTACCTTATGCTGAGGATGGGCTGCTAATGG	3240	4261	GGGCTGTGGGGGCGGCTTACGACATCATATGCTGATGATGCTCATGCTGCTGCTGCTGCT	4320
Db	3181	GTGAGCTGGCACTTATCATATGACACCTTACCTTATGCTGAGGATGGGCTGCTAATGG	3240	4261	GGGCTGTGGGGGCGGCTTACGACATCATATGCTGATGATGCTCATGCTGCTGCTGCTGCT	4320
Qy	3241	CCTCGGGAGCTTGGCGGCTGAGCTGCTTGGGGGATATCTTACGCTGAGGATGGGCTGCTAATGG	3300	4321	TACACCATCTCTTGGCATCGGAACTGCTTGTATCAAGCAGAGACAGCTGGGGTCTAGACT	4380
Db	3241	CCTCGGGAGCTTGGCGGCTGAGCTGCTTGGGGGATATCTTACGCTGAGGATGGGCTGCTAATGG	3300	4321	TACACCATCTCTCGGCATCGGAACTGCTTGTATCAAGCAGAGACAGCTGGGGTCTAGACT	4380
Qy	3301	CATTGTCTGGGAGCGGAGACAGCTGCTTGTGGGACATTTTACAGGACTTCCCGTCTC	3360	4381	AACGTGTGCTGCTTACAGCTACGCGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4440
Db	3301	CATGCTCTGGGAGCGGAGACAGCTGCTTGGGGGATATCTTACAGGACTTCCCGTCTC	3360	4381	AACGTGTGCTGCTTACAGCTACGCGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4440
Qy	3361	CGCCGAGCTTGGCGGGAGGCTCTCTTGGCCAGCTGATGGCTATACCTCCAGGGGCTG	3420	4441	GGAGTGGGCTTGGGAGGAGGCGGAGATCCCTTCTTATGGGAGGCGGATTCCTCTGTC	4500
Db	3361	CGCCGAGCTTGGCGGGAGGCTCTCTTGGCCAGCTGATGGCTATACCTCCAGGGGCTG	3420	4441	GGAGTGGGCTTGGGAGGAGGCGGAGATCCCTTCTTATGGGAGGCGGATTCCTCTGTC	4500
Qy	3421	GAGTCTTCTCGCCCGCATCACTGCTTACGCGGCTGAGCTGCTTACCTCCAGGACCAT	3480	4501	TTACATCAAGGGAGGAGACATCTGATCTTCTGCCATTCAAAGAAAAGTGTGACGAGCT	4560
Db	3421	GAGTCTTCTCGCCCGCATCACTGCTTATGCGGAGACAGCGGCTTGTGGGACCAT	3480	4501	ATACATCAAGGGAGGAGACATCTGATCTTCTGCCATTCAAAGAAAAGTGTGACGAGCT	4560
Qy	3481	AGTGTGAGCATGACGGGGCGGACAGACAGAAACAGCTGGGGAATTCAGGCTCTGTC	3540	4561	CGCGGCGGCTTTCGGGGTATGGGCTTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4620
Db	3481	AGTGTGAGCATGACGGGGCGGACAGACAGAAACAGCTGGGGAATTCAGGCTCTGTC	3540	4561	CGCGGCGGCTTTCGGGGTATGGGCTTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4620
Qy	3541	CACAGTCACTGCTCTTCTCGGAAACATTCATCTCGGGGCTTGTGGGCTGTCTACCA	3600	4621	CTCCGTAATACCAACTCAGGAGAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4680
Db	3541	CACGCTCACTGCTCTTCTCGGAAACATTCATCTCGGGGCTTGTGGGCTGTCTACCA	3600	4621	CTCCGTAATACCAACTCAGGAGAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4680
Qy	3601	TGGAGCTGGCAACAGACTTGGCGGCTCACGGGCTCCGCTCAGCAGATGTACTCCAG	3660	4681	GTATATCGGGGACTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4740
Db	3601	TGGAGCTGGCAACAGACTTGGCGGCTCACGGGCTCCGCTCAGCAGATGTACTCCAG	3660	4681	GTATATCGGGGACTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4740
Qy	3661	TGCTGAGGGGACTTGTAGGGGTGGCCAGCGGCTGGGACTAAATCTTTGGAGCGGCTG	3720	4741	CTTCAGTTTACAGCCCATTCACCATTAACCAACAGACTGTCCCTCAAGACGCTGTCTC	4800
Db	3661	TGCTGAGGGGACTTGTAGGGGTGGCCAGCGGCTGGGACTAAATCTTTGGAGCGGCTG	3720	4741	CTTCAGTTTACAGCCCATTCACCATTAACCAACAGACTGTCCCTCAAGACGCTGTCTC	4800
Qy	3721	CACGTGTGGAGCGGTGACCTGTATCTGCTCAGCGGAAACGCTGATGTCTATCCCGGCTG	3780	4801	ACGTAGCCAGCGGCGGCGGCGGAGAGCTGGGCTATTTATAGGTATGTGTTTC	4860
Db	3721	CACGTGTGGAGCGGTGACCTGTATCTGCTCAGCGGAAACGCTGATGTCTATCCCGGCTG	3780	4801	ACGTAGCCAGCGGCGGCGGCGGAGAGCTGGGCTATTTATAGGTATGTGTTTC	4860
				4861	CACGTGTGGAGCGGCTCAGGAAATGTTTGTAGTGTAGTGTCTCTGTGAGTGTGCTACGAGCG	4920

481 CGGCAAGGAAGACTTCGAGCGGTCCAGCGACAGTGGAGCGCGCAGCCCATCCCTAA 540
481 DB
541 AGATCGGCGCTCCACTGGCAAACTCTGGGAAACACAGATACCCCTGGCGCCCTATACGG 600
541 QY
541 GATCGGCGCTCCACTGGCAAACTCTGGGAAACACAGATACCCCTGGCGCCCTATACGG 600
541 DB
601 GAATGAGGAGCTCGGCTGGGAGAGATGGCTCTCTGTCCTCCCGAGGTTCCCGTCCCTCTTG 660
601 QY
601 GAATGAGGAGCTCGGCTGGGAGAGATGGCTCTCTGTCCTCCCGAGGTTCCCGTCCCTCTTG 660
601 DB
661 GGGCCCCAATGACCCCGGAGATAGGTGCGCGAAGCTGGGTAAGGTGATGATACCCCTAAC 720
661 QY
661 GGGCCCCAATGACCCCGGAGATAGGTGCGCGAAGCTGGGTAAGGTGATGATACCCCTAAC 720
661 DB
721 GTGGGCTTTGGCGACCTCATGGGATACATCCCTGTCGTGGGCGCCCGCTCGGCGCGT 780
721 QY
721 GTGGGCTTTGGCGACCTCATGGGATACATCCCTGTCGTGGGCGCCCGCTCGGCGCGT 780
721 DB
781 CGCCAGAGCTCTCGCGATCGGCTGAGAGTCTCTGGAGGACGGGTTAAATTTGCAACAGG 840
781 QY
781 CGCCAGAGCTCTCGCGATCGGCTGAGAGTCTCTGGAGGACGGGTTAAATTTGCAACAGG 840
781 DB
841 GAATTAACCGGTGCTCTTTCTATCTCTGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
841 QY
841 GAATTAACCGGTGCTCTTTCTATCTCTGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
841 DB
901 CCGGCTCTCGGCTCGGAGTGAAGACATCAGTACCGGCTACATGGTGAATACAGCTG 960
901 QY
901 CCGGCTCTCGGCTCGGAGTGAAGACATCAGTACCGGCTACATGGTGAATACAGCTG 960
901 DB
961 CACCAATGACAGATTAACCTGGAGCTTCCAGGCTGCTGCTCCAGCTCCCGGCTGCTGCT 1020
961 QY
961 CACCAATGACAGATTAACCTGGAGCTTCCAGGCTGCTGCTCCAGCTCCCGGCTGCTGCT 1020
961 DB
1021 CCGGCTCGGAAAGTGGGAAATCATCTCAGTCTGATACCGGCTCTCCAGGCTGCTGCTGCT 1080
1021 QY
1021 CCGGCTCGGAAAGTGGGAAATCATCTCAGTCTGATACCGGCTCTCCAGGCTGCTGCTGCT 1080
1021 DB
1081 GGTGACGCGCGCGGCGCTTACGAGGCTTTCGAGCGGCTTTCGAGCGGCTTTCGAGCGG 1140
1081 QY
1081 GGTGACGCGCGCGGCGCTTACGAGGCTTTCGAGCGGCTTTCGAGCGGCTTTCGAGCGG 1140
1081 DB
1141 GTCCGCGACGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
1141 QY
1141 GTCCGCGACGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
1141 DB
1201 AGCCCAAAATGTTTCAATGTTCTGCGCGCAGCACCACTGGTGTGTCAGAGCTGCAATGCTC 1260
1201 QY
1201 AGCCCAAAATGTTTCAATGTTCTGCGCGCAGCACCACTGGTGTGTCAGAGCTGCAATGCTC 1260
1201 DB
1261 CATCTACCTGGTACCATCATCTGACACCGCATGGGATGGGATGGGATGGGATGGGATGG 1320
1261 QY
1261 CATCTACCTGGTACCATCATCTGACACCGCATGGGATGGGATGGGATGGGATGGGATGG 1320
1261 DB
1321 GCCCAGGCTTACCATGATCTTGGCGTACGATGCGTGTCCCGAGGCTATATAGACAT 1380
1321 QY
1321 GCCCAGGCTTACCATGATCTTGGCGTACGATGCGTGTCCCGAGGCTATATAGACAT 1380
1321 DB
1381 CATTAAGCGGCTCATTTGGGCGTCTATGTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1440
1381 QY
1381 CATTAAGCGGCTCATTTGGGCGTCTATGTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1440
1381 DB
1441 GTGGGAAAGTGGTGTATCTCTGTTGGCGCGGGGTGGAGCGGCGGACCCATAC 1500
1441 QY
1441 GTGGGAAAGTGGTGTATCTCTGTTGGCGCGGGGTGGAGCGGCGGACCCATAC 1500
1441 DB
1501 TGTGGGGTTCGCGCGCAGACACCGGCGCTCACAGCTTATTTGATGGGCGC 1560
1501 QY
1501 TGTGGGGTTCGCGCGCAGACACCGGCGCTCACAGCTTATTTGATGGGCGC 1560
1501 DB
1561 CAGGCAAGAAATCCAGCTCGTTAAACCAATGGCAGCTGGGACATCAACCGCAGCCCT 1620
1561 QY

1561 CAGGCAAGAAATCCAGCTCATCAACCAACCAATGGCAGTTGGCACAATCAACCGCAGCCCT 1620
1561 DB
1621 GAATGCAATGACTCTCTGCAACACCGGCTTTATCGGCTCTCTGTTCTACACCAACAGCTT 1680
1621 QY
1621 GAATGCAATGACTCTCTGCAACACCGGCTTTATCGGCTCTCTGTTCTACACCAACAGCTT 1680
1621 DB
1681 CAACTCGTCAAGATGTCGGAACCGATGTCGCGCTCCCGCAGTATCGAGGCTTCCGGGT 1740
1681 QY
1681 CAACTCGTCAAGATGTCGGAACCGATGTCGCGCTCCCGCAGTATCGAGGCTTCCGGGT 1740
1681 DB
1741 GGGATGGGCGCTTGCATATGAGGATATGACCAATGACCAATCAGAGGATATGAGACCGTA 1800
1741 QY
1741 GGGATGGGCGCTTGCATATGAGGATATGACCAATGACCAATCAGAGGATATGAGACCGTA 1800
1741 DB
1801 TTGCTGGCACTACCCACCAAGGAGTGTGGCGTGTCTCCGGAAGACTGTGTGTGGCCC 1860
1801 QY
1801 TTGCTGGCACTACCCACCAAGGAGTGTGGCGTGTCTCCGGAAGACTGTGTGTGGCCC 1860
1801 DB
1861 AGTGTACTGTTTCAACCCCGCAGTGTAGTGGGACGACCGCAGGCTTGGAGCGCC 1920
1861 QY
1861 AGTGTACTGTTTCAACCCCGCAGTGTAGTGGGACGACCGCAGGCTTGGAGCGCC 1920
1861 DB
1921 CACTTACACGCTGGGAGGAAATGAGACAGATGCTTCTTATTTGAACAGACACTCGACACC 1980
1921 QY
1921 CACTTACACGCTGGGAGGAAATGAGACAGATGCTTCTTATTTGAACAGACACTCGACACC 1980
1921 DB
1981 GCTGGGCTCATGTTGCGCTGCAAGTGAATCTTCTTGGCTACCAAGACTTGGCG 2040
1981 QY
1981 GCTGGGCTCATGTTGCGCTGCAAGTGAATCTTCTTGGCTACCAAGACTTGGCG 2040
1981 DB
2041 CGCACACCTCTCGCTACTAGAGCTGACTTCAACCGCAGCAGGACTCTGTGTGGCCCCAC 2100
2041 QY
2041 CGCACACCTCTCGCTACTAGAGCTGACTTCAACCGCAGCAGGACTCTGTGTGGCCCCAC 2100
2041 DB
2101 GGAATGTTTGAAGAGCTCTGATACCACTTACCTCAATCGGCTCTGGGCTTGGCT 2160
2101 QY
2101 GGAATGTTTGAAGAGCTCTGATACCACTTACCTCAATCGGCTCTGGGCTTGGCT 2160
2101 DB
2161 CAGGCAAGTGGCTGATGACTACCCCTAGAGCTCTGGCATTTACCTCCCTGACAGTTAA 2220
2161 QY
2161 CAGGCAAGTGGCTGATGACTACCCCTAGAGCTCTGGCATTTACCTCCCTGACAGTTAA 2220
2161 DB
2221 CTATACCATCTTCAAAATAAGGATGTATGTGGGAGGGTGTAGCAGGCTCACGGCTGC 2280
2221 QY
2221 CTATACCATCTTCAAAATAAGGATGTATGTGGGAGGGTGTAGCAGGCTCACGGCTGC 2280
2221 DB
2281 ATGCAATTTCACTCGTGGGATGCTGCAACTTGGAGGACAGAGAGTCAACTGTC 2340
2281 QY
2281 ATGCAATTTCACTCGTGGGATGCTGCAACTTGGAGGACAGAGAGTCAACTGTC 2340
2281 DB
2341 TCCTTTGTCGACTCCACCGAATGGCCATTTTACCTGCTTACTCGGACTGCTGC 2400
2341 QY
2341 TCCTTTGTCGACTCCACCGAATGGCCATTTTACCTGCTTACTCGGACTGCTGC 2400
2341 DB
2401 GCGCTTGTGCACTGCTCTCCACTCCACCAAAACATCGTGGAGTCAATTCATGTA 2460
2401 QY
2401 GCGCTTGTGCACTGCTCTCCACTCCACCAAAACATCGTGGAGTCAATTCATGTA 2460
2401 DB
2461 TGGCTTATCACTGCTCCCTCACAAAATACATCGTCCGATGGGAGTGGGTAGTCTTAT 2520
2461 QY
2461 TGGCTTATCACTGCTCCCTCACAAAATACATCGTCCGATGGGAGTGGGTAGTCTTAT 2520
2461 DB
2521 CCGTCTTATAGCGGACCGGCTTGGCGCTGCTTATGGATGCTCATCTTGTGGGCA 2580
2521 QY
2521 CCGTCTTATAGCGGACCGGCTTGGCGCTGCTTATGGATGCTCATCTTGTGGGCA 2580
2521 DB
2581 GGGCCAAAGCAGCACTAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2640
2581 QY
2581 GGGCCAAAGCAGCACTAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2640
2581 DB
2641 TGGCTTCTATATTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2700
2641 QY
2641 TGGCTTCTATATTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2700
2641 DB

```

8701 GUAUUCUGCCUCCUGGUGACCCGCCACACGGAUAUAGCUGGAGCUGAUAACAUC 8760
8761 TTGCTCTCAATGTGTCTGTGGCGTGGCGCCACAGGCGCGGAGTACTACCTGAC 8820
8761 UUGCUCUCAAUUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 8820
8821 CAGAGACCCCTACCACTCAATCCGCGCGGCTGGGAAACAGTGTAGACACTCCCTGT 8880
8821 CAGAGACCCUACCAUCCAUCCGCGCGGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 8880
8881 CAUUAUGCUAGGAACAUAUCCAGUAGCGCCCAACCAUAGUGGUGUGUGUGUGUGUG 8940
8941 GATGACACACTTCTTCCATTCTCTGCTCAAGACAGCTGGACGACGACCTCACTT 9000
8941 GAUGACACAUUCUCCAUUCUCCAUUGGCGGCAAGAUUACUGGACGACGACCTCA 9000
9001 TGAGATGACGAGCGGTGACTCTCGGTGAGTCCCTTGGACCTCCAGCTATAATTGA 9060
9001 UGAGUAGUAGGAGCGUGUACUCCGUGAGUCCUUGGACUCCGACCAUUAUAGNA 9060
9061 GTTACATGCGCTTGACCTTTTCTCTGCAACATACACTCCCAAGACTGACGCGGT 9120
9061 GUUACACGGGUGUGAGCUUUCUCUGACACAUACACUCCCAAGACGACGCGGU 9120
9121 GGCTTACGCTCAGAAACTTGGGCGCCACCCCTCAGACGCTGGAGAGCGGCGCAC 9180
9121 GGUUACGCGCUCAGAAACTTGGGCGCCACCCCTCAGACGCTGGAGAGCGGCGCAC 9180
9181 TGAGTACGCGCGCTCTATCTCCGCGGAGAGCGCGCTTGGCGTGCATATCT 9240
9181 UGAGUAGGCGGUGUCCUACUCCGUGGCGGAGAGCGCGGUGUGUGUGUGUGUGUG 9240
9241 CTTCAATGCGCGGTGAAGCAAGCTCAAACTCACTCAATGCGGAAAGCGCGCTCT 9300
9241 CUUACACUGGCGGUGAAGCAAGCTCAAACTCACTCAATGCGGAAAGCGCGCTCT 9300
9301 GGATTTATCAGCTGTTCCAGCTCGCGGCGGCGGCGGCGGACATTTATCACAGCGT 9360
9301 GGAUUAUACGUGGUGUACUGUGCGCGCGGCGGCGGCGGACAUUAUACAGCGUG 9360
9361 GCGTGCGGACCGCGCTATTGCTCTTGGCTACTCTACTTTTGTAGGGGTAGGCT 9420
9361 GCGUGCGGACCGCGCGUUAUUCUUGGCGUACUCCUUAUUAUUAUUAUUAUUAU 9420
9421 TTTCCTACTCCCGCTCGGTAGAGCGGACACATTAAGTACTCACTCACTCACTGTC 9480
9421 UUUUUAUACCGCGUGUAGAGCGGACACAUUAUUAUUAUUAUUAUUAUUAUUAU 9480
9481 CTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 9540
9481 CUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU 9540
9541 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 9588
9541 UUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU 9588

```

RESULT 2
 US-07-925-695-2
 ; Sequence 2, Application US/07925695
 ; Patent No. 5428145
 ; GENERAL INFORMATION:
 ; APPLICANT: OKAMOTO, Hiroaki
 ; APPLICANT: NAKAMURA, Tetsuo
 ; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOME,
 ; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, ANTIGEN, ANTIBODY AND
 ; TITLE OF INVENTION: DETECTION SYSTEMS
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Beveridge, DeGrandi, Weillacher & Young

```

; STREET: 1850 M Street, N.W., Suite 800
; CITY: Washington
; STATE: D.C.
; COUNTRY: US
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/925,695
; FILING DATE: 19920807
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 287402/91
; FILING DATE: 09-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 360441/91
; FILING DATE: 05-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Weillacher, Robert G.
; REGISTRATION NUMBER: 20,531
; REFERENCE/DOCKET NUMBER: 06/87-48009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2811
; TELEFAX: (202) 659-1462
; TELEX: WUI 64470
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9589 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-07-925-695-2

```

```

Query Match 92.4%; Score 8970.4; DB 1; Length 9589;
Best Local Similarity 96.0%; Pred. No. 0;
Matches 9202; Conservative 0; Mismatches 386; Indels 0; Gaps 0;

QY 1 ACCCGCCCTAATAGGGGCGACACTCCGCGCATGAATCACTCCCTCTGAGGAACACTACTGT 60
DB 1 ACCCGCCCTAATAGGGGCGACACTCCGCGCATGAATCACTCCCTCTGAGGAACACTACTGT 60
QY 61 CTTACGCGAAGAGCGCTTAGCCATGCGCGTGTAGTATGAGTGTCTACAGCCTCCAGGCC 120
DB 61 CTTACGCGAAGAGCGCTTAGCCATGCGCGTGTAGTATGAGTGTCTACAGCCTCCAGGCC 120
QY 121 CCCCCCTCCCGGAGAGCCATAGTGGTCTCGGAAACCGGTGAGTACACCGGAATTGCGGG 180
DB 121 CCCCCCTCCCGGAGAGCCATAGTGGTCTCGGAAACCGGTGAGTACACCGGAATTGCGGG 180
QY 181 AAGACTGGGTCTCTTCTTGATTAACCCACTCTATGCGCGGCGCATTTGGCGGTGCGCCCG 240
DB 181 AAGACTGGGTCTCTTCTTGATTAACCCACTCTATGCGCGGCGCATTTGGCGGTGCGCCCG 240
QY 241 CAGAGCTGTAGCGGAGTAGCGTGTGGTTCGGAAGCGCTTGTGTACTGCTGTATAGGG 300
DB 241 CAGAGCTGTAGCGGAGTAGCGTGTGGTTCGGAAGCGCTTGTGTACTGCTGTATAGGG 300
QY 301 TGCTTGGAGTGCCTCCCGGAGGTCTCGTAGACCGTGCATCCATGAGCAAACTCTAAACC 360
DB 301 TGCTTGGAGTGCCTCCCGGAGGTCTCGTAGACCGTGCATCCATGAGCAAACTCTAAACC 360
QY 361 TCAAGAAACCAAAAGAAACACCAACCGTCCGCCACAGAGCTTAAGTTTCCGGCGG 420
DB 361 TCAAGAAACCAAAAGAAACACCAACCGTCCGCCACAGAGCTTAAGTTTCCGGCGG 420
QY 421 CGGCCAGATGTTGGCGGAGTATCTTGTTCGCGCGCAGGCGGCCAGGTGTGGTGTGCG 480
DB 421 CGGCCAGATGTTGGCGGAGTATCTTGTTCGCGCGCAGGCGGCCAGGTGTGGTGTGCG 480
QY 481 CGGCACAGGAAGACTTTCGAGCGGTCCAGCCACGTTGGAAAGCGCGCGCCATCCCTAA 540

```

7621	GTCTCGTGTACTTGTCTCCGAGGAGCAGCTCGTCTGTGTCTGTCTCAATGTCACTC	7680
7621	GUCUUGGUCUACUUGCCUGCGAGGAGCAGTCCUGGUGUGUGUGUGUCAUACUC	7680
7681	CTGAGCCGGGCTCTAATACTCTTGTAGTCCCGAGAGGAGAGTGTACCGATTAAACC	7740
7681	CUGGACCCGGGUCUUAUACUUCUUGUAGCCCGAGAGGAAAGUUGCCAUUUGGCC	7740
7741	CTTGAGCAACTCCCTGTTCGATATCAAAACAAGGTGTACTGTACCAAAACAAGAGCGC	7800
7741	CUUGAGCAACUCCUGUUGCGAUACAACAAGGUGUACUGUACCAUCAACAAGAGCGC	7800
7801	CTCACTAAGGCTAAAAAGTAACTTTTGTATGATGCAAGTGTCTGACTCTCTACTACA	7860
7801	CUCAUUAAGGCUAAAAGGUAAAUUUUGUAGGAUGCAAGCGCUGCAUUAUGA	7860
7861	CTCAGTCTTAAGGACATTAAAGCTAGCGGCTCCAAAGTCAACGAGGCTCCTCAACAT	7920
7861	CUCAGUCUUGAAGGACAUUUAAGCUGAGCGCCUCCAAGGUCACCGCAAGGCUUCACUUU	7920
7921	GGAGGAGCTTGCAGATTAAACCCACCCCATCTTGCAGAGTCTAAATATGGTTTGGGCG	7980
7921	AGAGGAGCCUGCCAGUUUAUCCACCCCAUCUUGCAAGUCCAAGUAGGUUUGGCGC	7980
7981	TAAAGGTCGCGAGCTTGTTCGGGAGGCGGTAAACCACTAAGTCCGTGTGGAAGGA	8040
7981	UAGAGGUCGCGAGCUUUGCCGAGAGCCGUUAACCAUCAAGUCCGUGUGGAAGGA	8040
8041	CCTCTGAGAGACTCAGAAACACCAATTTCCCAACCACTTATGSCCAAAAATGAGTGTT	8100
8041	CCUCUGGAAGACAACAACAACCAUUCUUAACAACAUGGCCAAAAAUGAGGUGUU	8100
8101	CTCGTGTGACCCCAAGAGGGGGCAAGAAAGAGCTCGCTCTATCTGTATACCTGACCT	8160
8101	CUGGUGAGCCCAACCAAGGGGGUAGAAGAGCAGCUCGCCUUUACCUUACCCUGACCU	8160
8161	CGCGTCAAGGTCTGCGAGAAAGATGGCCCTTTTATGACATTACAAAAACTTCTCAGGC	8220
8161	CGCGUCAGGUCUGCGAGAAAUUGGCCUUUAUGAUUAUCAACAAAAAGCUCCACAGC	8220
8221	GGTATGGGGCTTCTTATGATTTCCAGTATTTCCCGCTCAGCGGGTAGAGTTTCTCTT	8280
8221	GGUAGGGGGCUUUAUGGAUUCAGUACUCCCGCUCAGCGGGUGAGUUUCUUU	8280
8281	GAAGCATGGCGGAAAGAGGACCTCTATCGGTTTTTCTGATGATACCCGATGCTTGA	8340
8281	GAAGCUGGGCGGAAAGAAAGAACCCUUGGGUUUUUGUAGUACCCGAGUUGCUGA	8340
8341	CTCAACCTCACTAGAGAGAGCATCAGGACTGAGGAGTCCATATATCGGGCTGTCTCTT	8400
8341	CUCAACCGUACUGAGAGAGCAUCAGGACUGAGGAGUCCAUAUAUCGGGCUUGUCCUU	8400
8401	GCCGAGAGGCCCACTGCCATACACTCGCTAACTGAGAGACTTTTACGTGGAGAGGCC	8460
8401	GCCGAGAGGCCCAACAUGCAUACAUCUACUGACUGAGAGACUUUACUGGGAGGCC	8460
8461	TATGTTCAACAGCAGGSCCAAACTCGGGGTACAGGCGTTCGCGCCAGCGGGTCT	8520
8461	CAUGUUCAACAGCAAGGGCCAGACUUGCGGUAACAGGGUUGCCGCGCAGCGGGUCU	8520
8521	CACCACTAGCATGGGGAACACCATCACAATGCTAGCTGAAAGCCTTAGCGGCTGTAAAGC	8580
8521	UACCACUAGCAUGGGGAACACCAUCAUGCUUUGUAAAGCCUUAAGCGGCCUUAAGGC	8580
8581	TGCAAGGATATCGGCCCACAATCTGGTATGCGGATGACTTGGTTGTCTCATCTCAGA	8640
8581	UGCAGGGAUAAUUGCGGCCAACAUUGUGGUUAGCGGCGUGACUUGUUGUCAUCUCAGA	8640
8641	AAGCCAGGACCGAGGAGCAGCGGAACCTCGAGGCCCTTACCGAGGCTATGACACAG	8700
8641	GAGCCAGGGACCGAGAGGACGAGCGGAACUUGAGAGCCUUAUCGAGGCUUAGACCGAG	8700
8701	GTAATTCGCCCTCTGTGTGACCCCGCCAGACCCGAGATGATGATCTGGAGCTGATAACATC	8760

2161 CAGCCAAAGTGCTGATGACCTACCCCTACAGGCTCTGGCAATTACCCCTGACAGTTAA 2220
2161 CAGCCAAAGTGCTGATGACCTACCCCTACAGGCTCTGGCAATTACCCCTGACAGTTAA 2220
2221 CTATACATCTTCAAAATAGAGTATGATGGAGGGTTGAGCAGCAGGCTCACGCTGC 2280
2221 CUAACCAUCUCAAUAAGGAUGAUGUGGGGGGUGGAGCAGCAGGCTCACGCTGC 2280
2281 ATGCAATTTCACTGCTGGGATCGTTGCACTTTGGAGGACAGAGAGAGTCAACTGTC 2340
2281 GUGCAAUUUCAUCUGUGGGGAUGUUGCAUUGGAGGACAGAGAGAGTCAACTGTC 2340
2341 TCCCTTTGTCACCTCCACACAGGAATGGGCCATTTTACCTTGTCTTACTCGACCTGCC 2400
2341 UCCUUUGUGCAUCCACACAGGAGUGGGCCAUUUUACCUUGCAUUAUCGAGCCUGCC 2400
2401 CGCCTTGTGCACTGCTTCTCCACCTCCACCAAAACATGCTGGAAGTCAATTCATGTA 2460
2401 CGCCTTGTGCACTGCTTCTCCACCTCCACCAAAACATGCTGGAAGTCAATTCATGTA 2460
2461 TGGCTATACCTGCTCCCTACAAATACATCGTCCGATGGAGTGGTAATCTTTATT 2520
2461 UGGCCUAUACCTGCTCCCTACAAATACATCGTCCGATGGAGTGGTAATCTTTATT 2520
2521 CTTGCTCTTACGCGAGCCAGGGTTTGGCTCTGCTTATGATGCTCATCTTGTGGGCCA 2580
2521 CCUGCUUAGCGGACCGCAGGUGUUGGCGUUAUGGAGUUGCUUUGGCGCA 2580
2581 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2640
2581 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2640
2641 TGGCTCTTATTTTGTCTTCTTTTGTGCTGCTTGTGATCATCAAGGCTCGGTAGT 2700
2641 UGGCUUCCUAUACUUGUUGCAUUCUUGGCGUGUGGUAUCAUAAAGGUGGCGGAGU 2700
2701 CCCCTTGTGCTTATTTTGTCTTCTTCTGCTGCTTCTTGTGCTTCTGCTTCTGCTTCT 2760
2701 CCCUUGGCUUUAUUCUUGGCGUGUGGCGUUAUGGAGUUGCUUUGGCGCA 2760
2761 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2820
2761 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2820
2821 GGTAAATGATCACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2880
2821 GGUACGUAUCUUCUUGGCGUGUGGCGUUAUGGAGUUGCUUUGGCGCA 2880
2881 GTGGTGTGCTTATTTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2940
2881 GUGGUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGU 2940
2941 TATGCAAGTGTGCGGCTGCTGATGCTATATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 3000
2941 UAUGCAGGUGCGGUGGCGUGGAGGUAUUGGCGUGUGGCGUUAUUGGCGG 3000
3001 TGTGTTGTGATCAATTAACAGTGGCTTGTGGGCTGCTTGTGGGCTGCTTCTTCTTCTTCT 3060
3001 UGUGUGUUGUUGAUAUAAACCAAGGUGGUGUUGGCGUGUGGCGUUAUUGGCGG 3060
3061 AGGTGCTTGTGACCGGCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3120
3061 AGGUGUUGUUGGCGGUGGCGUUAUUGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3120
3121 CATGGCAAGGATCTCGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3180
3121 CAUGGUAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3180
3181 GTGGATGGGATCTTATATGACCACTTATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3240
3181 GUGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3240

3241 CCTGGGAGACCTGGGGTGGCGCTTGGAGCTTATCATCTTCTGCTGCTGCTGCTGCTGCTGCT 3300
3241 CUUGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3300
3301 CATTTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3360
3301 CAUCGUGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3360
3361 GCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3420
3361 GCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3420
3421 GAGTCTTCTGCGCGCGCGCTTATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3480
3421 GAGUUCUUCG 3480
3481 AGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3540
3481 AGUGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3540
3541 CACAGTCTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3600
3541 CACGUGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3600
3601 TGGAGTCTGCGCGCGCGCTTCTGCGCGCGCGCTTCTGCGCGCGCGCTTCTGCGCGCGCGCT 3660
3601 UGGAG 3660
3661 TGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3720
3661 UGUGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3720
3721 CACGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3780
3721 CACGUGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3780
3781 AAGAG 3840
3781 AAGAG 3840
3841 GTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3900
3841 GUUCGUGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3900
3901 TGTGCTCTCTGCGGCGGCTTAAAGTCCATAGATTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3960
3901 CGUGUGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3960
3961 CGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4020
3961 CGUACGUGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4020
4021 TCAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4080
4021 UCAGGUGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4080
4081 ATATCTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4140
4081 GUAGGUGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4140
4141 GTTGGGAG 4200
4141 GUUUGGAG 4200
4201 GACTGTGAG 4260
4201 GACUGAG 4260
4261 GGGCTGTGAG 4320
4261 GGGGUGGAG 4320
4321 TACCACCATCTTCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4380

Best Local Similarity 74.6%; Pred. No. 0;
Matches 7151; Conservative 2051; Mismatches 386; Indels 0; Gaps 0;

1	ACC	CGC	CCCTTA	TAG	GGG	CGG	CAC	TCT	CGG	CCATGA	TACT	CTCCCTGTG	AGG	AACTACT	GT	60
1	ACC	CGC	CCCTTA	AUA	AGG	GGG	CGG	CAC	TCT	CGG	CCATGA	TACT	CTCCCTGTG	AGG	AACTACT	GT
61	CTT	CAC	GCAG	AAAG	CGT	CTAG	CCAT	TGG	CGG	TTAG	TAT	GAG	TG	TG	TAG	120
61	CTT	CAC	GCAG	AAAG	CGT	CTAG	CCAT	TGG	CGG	TTAG	TAT	GAG	TG	TG	TAG	
121	CCCC	CTCC	GGG	GAG	CCAT	ATG	TGT	CTC	GG	AA	CCGG	TGAG	TAC	CAC	CGGA	180
121	CCCC	CTCC	GGG	GAG	CCAT	ATG	TGT	CTC	GG	AA	CCGG	TGAG	TAC	CAC	CGGA	
181	AAG	ACT	GGT	CTCT	TTCT	TGG	ATA	AA	CC	CA	TCT	ATG	TG	CC	CGG	240
181	AAG	ACT	GGT	CTCT	TTCT	TGG	ATA	AA	CC	CA	TCT	ATG	TG	CC	CGG	
241	CAAG	ACT	GCT	AGC	AGT	AGC	TGT	GG	TTG	CG	AA	AG	CGC	TCT	TG	300
241	CAAG	ACT	GCT	AGC	AGT	AGC	TGT	GG	TTG	CG	AA	AG	CGC	TCT	TG	
301	TGCT	TTG	CAG	ATG	CCCC	GGG	AGG	TCT	CGT	AGA	CCGT	GCA	CA	ATG	AG	360
301	TGCT	TTG	CAG	ATG	CCCC	GGG	AGG	TCT	CGT	AGA	CCGT	GCA	CA	ATG	AG	
361	TCA	AG	AAAA	AA	CA	CA	AA	AA	CA	CA	AA	CGT	CG	CC	CA	420
361	TCA	AG	AAAA	AA	CA	CA	AA	AA	CA	CA	AA	CGT	CG	CC	CA	
421	CGG	CCAG	AT	CGT	TGG	CGG	AGT	AT	ACT	TGT	TG	CG	CG	CG	CG	480
421	CGG	CCAG	AT	CGT	TGG	CGG	AGT	AT	ACT	TGT	TG	CG	CG	CG	CG	
481	CGC	GA	CA	AG	GA	GA	GA	GA	GA	GA	GA	GA	GA	GA	GA	540
481	CGC	GA	CA	AG	GA	GA	GA	GA	GA	GA	GA	GA	GA	GA	GA	
541	AG	AT	CGG	CG	CT	CA	CT	GG	CA	AA	TCT	GG	GG	AA	AA	600
541	AG	AT	CGG	CG	CT	CA	CT	GG	CA	AA	TCT	GG	GG	AA	AA	
601	GA	AT	GAG	GAC	T	CGG	CT	GG	C	A	G	A	T	G	G	660
601	GA	AT	GAG	GAC	T	CGG	CT	GG	C	A	G	A	T	G	G	
661	GGG	CCC	CA	AT	G	A	T	A	C	CC	CG	G	A	T	A	720
661	GGG	CCC	CA	AT	G	A	T	A	C	CC	CG	G	A	T	A	
721	GTG	CG	CT	TTG	CG	CA	CT	CA	T	GGG	GT	A	CA	T	CC	780
721	GTG	CG	CT	TTG	CG	CA	CT	CA	T	GGG	GT	A	CA	T	CC	
781	CGC	CA	G	A	G	A	G	A	G	A	G	A	G	A	G	840
781	CGC	CA	G	A	G	A	G	A	G	A	G	A	G	A	G	
841	GAA	CTT	A	C	CG	T	T	G	T	C	T	T	T	C	T	900
841	GAA	CTT	A	C	CG	T	T	G	T	C	T	T	T	C	T	
901	CCG	GT	CT	CG	G	T	CC	GA	AG	T	GA	AG	A	CA	T	960
901	CCG	GT	CT	CG	G	T	CC	GA	AG	T	GA	AG	A	CA	T	
961	CAC	CA	AT	G	A	C	A	T	T	A	C	T	T	G	G	1020
961	CAC	CA	AT	G	A	C	A	T	T	A	C	T	T	G	G	
1021	CCG	CT	CG	CA	AA	AG	T	GG	GA	T	G	A	T	C	T	1080
1021	CCG	CT	CG	CA	AA	AG	T	GG	GA	T	G	A	T	C	T	

1021	CCGUGCGAAGAGUGGGAAUAUCAUCUGGUGUGGAUACCGGUGUACACCGAAUGUGGC	1080
1081	CGTGACGGGCCCGGCCCTCAGCAGAGGGCTTGGGACGCAATCGACATGCTGTGTGTAT	1140
	CGUGACAGCGCGCGCCUACACGACGGGCUUACGGACGCAUAUGACUGGUGUGUAU	1140
	CGUGACAGCGCGCGCCUACACGACGGGCUUACGGACGCAUAUGACUGGUGUGUAU	1140
1141	GTCCGCCACGCTCTGCTCTGCCCTCTACGTGGGGGACCTCTGCGGTGGGGTGATGCTCGC	1200
1141	GUCGCCACGCGUCGUCUCGCUUUUACUGGGGGACCCUCUGCGUGGGUGAUGCUGGC	1200
1201	AGCCCAATGTTCAATGTTCTCGCCCGCAGCACCACTGGTGTGTTCTCCAAAGACTGCAATTGCTC	1260
1201	AGCCCAGUUGUCAUGUCGCGCACAGCACCAUGGUUUGUGAAGACUGCAAUUGCUC	1260
1261	CATCTACCTGGTACCAATCACTGGACACCGCATGGCATGGGACATGATGTAATGCTGTC	1320
1261	CAUCUACCCUGUACCAUCAUGGACACCGCAUGCGUGGGGACAUAGAUGAAUCUGGUC	1320
1321	GCCACCGCTACCATGATCTTGGCGTACGGCATGCTGTCTCCCGAGGTCAATTATAGACAT	1380
1321	GCCACCGCUACCAUGAUCUGGCGUACGCGAUGCGGUCGCCGAGGUGUACUAUAGACAU	1380
1381	CATTAGCGGGCTCAATTGGGGCGTCATGTTGGGCTTGGCTTCTCTATGACGGGAGC	1440
1381	CAUUGCGGGGCUCAUUGGGGCUCAUGUUCGCUUACCUUUAUGACAGGAGC	1440
1441	GTGGGGAAGTCTGTTGTCATCTTCTGTTGGCGCGCGGGGTGGACGGCGCACCATATC	1500
1441	GUGGGCAAAAGUCUGUGUCAUUCUUUUGUGCGCGCGCGGGGUGGACGGCAACCAUAC	1500
1501	TGTTGGGGTCTGCGCGGACAGCACCGGGCGCTCACAGCTATTTTGACATCGGGCC	1560
1501	CGUUGGGGUGUUCACCGCGCAUACCGCAGACCCGCGCAUGUUCUCCUUGUGGC	1560
1561	CAGCAGAAATCCAGCTCGTTTAAACACCAATGGCAGCTGGCACATCAACCGCACCGCCT	1620
1561	CAGCAGANAUAUCAGCUCAUCAACCAUUGGAGUUGGACAUCAACCGCACCGCCU	1620
1621	GAACTGCAATGACTCTTGGCACACCGGCTTTATCGCGTCTCTGTTCTACACCAAGCTT	1680
1621	GAAUGCAAGUACUUCUUGACACCGGCUUCUCGCGUCACUUGUUCACACCCACAGCUU	1680
1681	CAACTGTCAGGATGTCGGAAACGATGTCGCGCTGCGCGAGTATCGAGGCTTCGGGT	1740
1681	CAACUGTGCAGGAUGUCCGGAACCAUGUCGCGCUGCGCGAGUACUGGAGCCUUGCGGU	1740
1741	GGATGGGGCCCTTGCATAATATGAGTAATGTCAACCAATCCAGAGGATATGAGACCTTA	1800
1741	GGGAUGGGGCCCUUACAAUAUAGGAGCAUUGUACCAUCCAGAGGAUAUGAGACCGUA	1800
1801	TTGTGCGACTPACCCACCAAGGCAGTGGCGTGTCTCCGGAAGACTGTGTGTGGCCC	1860
1801	UUGUGGCACUACCCACCAAGCAGUGUGUGUAGUCUCGCGAGCTCUCUGUGUGGCC	1860
1861	AGTGTACTGTTTACCCCGACCGCAGTGGTAGTGGGACGACCGCAGCTTGGAGCGCC	1920
1861	AGUGUACUGUUCACCCCGACCGCAGUAGUGGGUACGACCGAGUACUUGGAGCGCC	1920
1921	CACTTACAGTGGGGGAGAATGAGACAGATGTCTTCTTATTGAACAGCACTCGACCAAC	1980
1921	CACUUAACGUGGGGGAGAAUGAGACAAGUUCUUAUUGAACACGACACUUGACAC	1980
1981	GCTGGGCTCATGGTTCGCGCTGCACGTGGATGAATCTTTCTGGCTACACCAAGACTTGGG	2040
1981	GCAGGGGUCAUGGUGGUCGACCGUGGAUAAUCCACUGGCTUACACCAAGACUUGCG	2040
2041	CGCACACCTGCGGTACTAGAGTCGACTTCAAACCGCAGCAGGACCTGTTGTGCCCCAC	2100
2041	CGCACACCTGCGGTCAUUGAGGUCGACUUAUUGCCAGCAUGGCAUUGUUGGCCAC	2100
2101	GGACTGTTTTAGGAGCATCTCTGATACCATTTACCTCAAAATGCGGCTCTGGGCCCTGGCT	2160
2101	GGAUGUUUUAAGGAAGCAUCCUGAUACCAUUAUUAUUAUUGGUGGUCUUGGGCCUUGGU	2160

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 12, 2004, 06:47:41 ; Search time 470 Seconds
(without alignments)
9119.727 Million cell updates/sec

Title: US-09-980-559-1

Perfect score: 9711

Sequence: 1 accgcgcctaataggcg.....ggtctctgcgatcatgt 9711

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/1/ina/5A COMB.seq.*
- 2: /cgn2_6/prodata/1/ina/5B COMB.seq.*
- 3: /cgn2_6/prodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/prodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/prodata/1/ina/PCUS COMB.seq.*
- 6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8970.4	92.4	9589	1	US-07-925-695-1
2	8970.4	92.4	9589	1	US-07-925-695-2
3	6031.4	62.1	9511	1	US-07-925-695-7
4	6028.6	62.1	9511	1	US-07-925-695-6
5	4625.6	47.6	9646	3	US-08-811-566-1
6	4625.6	47.6	9646	4	US-09-034-756-1
7	4620	47.6	12980	3	US-08-811-566-5
8	4620	47.6	12980	4	US-09-034-756-5
9	4534.4	46.7	9599	3	US-09-014-416-6
10	4523.2	46.6	9599	3	US-09-014-416-7
11	4505.8	46.4	9595	3	US-09-014-416-4
12	4428.6	45.6	9401	5	PCT-US91-02225-9
13	4425.8	45.6	9401	1	US-07-910-760-9
14	4425.8	45.6	9401	1	US-08-440-519-9
15	4425.8	45.6	9401	4	US-08-440-549-9
16	4425.8	45.6	9401	4	US-08-823-895A-25
17	4421	45.5	9416	1	US-08-324-977-1
18	4421	45.5	9416	2	US-08-384-616-1
19	4421	45.5	9416	2	US-08-904-686A-1
20	4421	45.5	9416	3	US-09-315-850-1
21	4421	45.5	9416	4	US-08-823-895A-27
22	4413.2	45.4	9401	2	US-08-432-693-1
23	4413	45.4	9379	3	US-09-388-874-1
24	4413	45.4	9379	4	US-09-316-359-1
25	4413	45.4	9416	3	US-08-811-566-19
26	4413	45.4	9416	4	US-09-034-756-19
27	4412.6	45.4	9379	3	US-08-444-818-176

28	4410.8	45.4	9472	4	US-08-150-204E-96	Sequence 96, Appl
29	4392.2	45.2	9416	4	US-08-823-895A-26	Sequence 26, Appl
30	4311.4	44.4	9185	3	US-08-444-818-122	Sequence 122, App
31	4311.4	44.4	9185	3	US-08-444-818-123	Sequence 123, App
32	4106	42.3	9030	1	US-08-324-977-13	Sequence 13, Appl
33	4106	42.3	9030	2	US-08-384-616-13	Sequence 13, Appl
34	4106	42.3	9030	3	US-08-904-686A-13	Sequence 13, Appl
35	4106	42.3	9030	3	US-09-315-850-13	Sequence 13, Appl
36	4083.6	42.1	8987	3	US-08-444-818-137	Sequence 137, App
37	3723.6	38.3	3970	1	US-07-925-695-3	Sequence 3, Appli
38	3678.4	37.9	8316	3	US-08-444-818-88	Sequence 88, Appl
39	3514.6	36.2	7917	1	US-08-324-977-31	Sequence 31, Appl
40	3514.6	36.2	7917	2	US-08-384-616-31	Sequence 31, Appl
41	3514.6	36.2	7917	2	US-08-904-686A-31	Sequence 31, Appl
42	3514.6	36.2	7917	3	US-09-315-850-31	Sequence 31, Appl
43	3499.6	36.0	7863	1	US-08-324-977-35	Sequence 35, Appl
44	3499.6	36.0	7863	2	US-08-384-616-35	Sequence 35, Appl
45	3499.6	36.0	7863	2	US-08-904-686A-35	Sequence 35, Appl

ALIGNMENTS

RESULT 1

US-07-925-695-1
; Sequence 1, Application US/07925695
; Patent No. 5428145
; GENERAL INFORMATION:
; APPLICANT: OKAMOTO, Hiroaki
; APPLICANT: OKAMURA, Tetsuo
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOME,
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, ANTIGEN, ANTIBODY AND
; TITLE OF INVENTION: DETECTION SYSTEMS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Beveridge, DeGrandi, Weilacher & Young
; STREET: 1850 M Street, N.W., Suite 800
; CITY: Washington
; STATE: D.C.
; COUNTRY: US
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/925,695
; FILING DATE: 19920807
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 287402/91
; FILING DATE: 09-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 360441/91
; FILING DATE: 05-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Weilacher, Robert G.
; REGISTRATION NUMBER: 20,531
; REFERENCE/DOCKET NUMBER: 06/87-48009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2811
; TELEFAX: (202) 659-1462
; TELEX: WUI 64470
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9589 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
US-07-925-695-1

Query Match 92.4%; Score 8970.4; DB 1; Length 9589;

QY 6366 GCCTCCCTTTGTCTCTGTCAAAAGGGGTAAAGGGCGTGTGGCGGCGACTGGCATCA 6425
 Db 6355 GGATTCCTTTGTCTCTGTCAAGCGGGGTATAGGGGGGTCTGGCGAGGAGACGGCAITTA 6414
 QY 6426 TGACACACAGGTGTCTTGGGGGCCCAATATCTCTGGCAATGTCCGGTGGGTCCATGA 6485
 Db 6415 TGCACACTCGCTGCACTGTGGAGCTGAGATCACTGGACATGTCAAAAACGGGACGATGA 6474
 QY 6486 GAATACAGGGGCTTAAGACCTGATGAATATCTGGCAGGGGACTTCTCTATCAATTTGT 6545
 Db 6475 GGATCTGGTCTCTAGGACCTGAGGAACATGTGGAGTGGGAGCTTCCCATTAACGCTT 6534
 QY 6546 ACACGAGGGCCAGTGTGGTCCGAAACCCGCGCCAACTTTAAAGTGCAGCTCTGGAGGG 6605
 Db 6535 ACACGAGGGCCCTGTACT 6594
 QY 6606 TGGGGGCTCAGAGTACGGGAGGTGAGCGAGCAGCGAGGTCAACCACTACATTAACAGGAC 6665
 Db 6595 TGTCTGAGAGGAATAGTGGAGATAAGGCGGGTGGGGAGCTTCCACTAGTATCGGGTA 6654
 QY 6666 TCACCACTGATAACTTGAAGTCCCTGCGCAACTACCTCTCTCTCTCTCTCTCTCTCTCT 6725
 Db 6655 TGACTACTGACAACTTAAATGCGGTGCGAGATCCCATCGCCGAAATTTTTCACAGAAT 6714
 QY 6726 TGGACGAGTGCAGATCCATAGGTTTGGCCCCACACCGAGCGGTTTTTTCGGGATGAGG 6785
 Db 6715 TGGACGGGGTGGCTACATAGTGTGGCGCCCTTGGAGCCCTTGTCTGGGGAGGAGG 6774
 QY 6786 TCTCGTTCTCGGTTGGCTTAATTTCAATTTGTCTGGGGTCCAGCTTCTCTGGGACCTTG 6845
 Db 6775 TATCATTTAGAGTGGAGTCCAGAGTACCGGGTGGGGTGGCAATTAACCTTGGAGCCGG 6834
 QY 6846 AACCCGACAGAGTATGATGCTCCATGCTACAGATCCATCTCATATACAGCGCGAGA 6905
 Db 6835 AACCGAGTGGGTTGACGTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6894
 QY 6906 CTGACGCGGGCTTTAGCGGGGGTCAACCCCATCCGAGGCAAGCTCTCGGGCAGCC 6965
 Db 6895 CGCGGGGAGAGTTGGCGAGAGGTTCACCCCTTCTATGGCAGCTCTCGGGCAGCC 6954
 QY 6966 AGCTATCGGACCATCGCTGCGAGCACTGTGACCACTGACCAAGCAAGCTTATGATGG 7025
 Db 6955 AGCTGTCCGCTCCATCTCTCAAGGCACTTGGACCGCCCAACCATGACTCCCTCTGAC 7014
 QY 7026 ACATGTGGATGTAACCT-----GTTTCATGGGGCGATGCTCGATAG 7073
 Db 7015 AGCTATAGAGTAACTCTGTGAGGAGAGATGGCGCAACATCAACAGGGTTG 7074
 QY 7074 AGTCTGGGTCCAAAGTGGTCTGTGAGTCTCTCGACCAATGGTTCGAAGAAAGGAGCG 7133
 Db 7075 AGTCAGAGAACAAAGTGGTATCTGAGCTCTCTCGATCGCTTGTGGCAGAGGAGATG 7134
 QY 7134 ACCTGAGCCTTCGATACCAATCAGAAATACATGCTCCCAAGAAAGGTTCCCAACAGCTT 7193
 Db 7135 AGCGGAGGTCTCCGTACCGCGAGAAATCTGCGGAATGCTCGAGATTCGCGCGGGCCC 7194
 QY 7194 TACCGGCTGGGACCGGCTGATTAACCCACGCTTGTGGATCGTGGAAAGGCGAG 7253
 Db 7195 TGCCGCTTGGGCGCGCGGAGTCAACCCCGCTGTAGAGAGCTGGGAAAGCCTG 7254
 QY 7254 ATTACCAACCGGCGACTGTTGGGGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7313
 Db 7255 ACTACGAAACCACTGTGGTCCATGGTGGCGGTACCACTCAACGGTCCCTCTCTGTGC 7314
 QY 7314 CTCGCCAAGAGGCGCGGACAGTGGGCTTAAGTGAAGATCTCATAGAGATGCGCCTTC 7373
 Db 7315 CTCGCGCTCGGAAAGAGCGTACGGTGTCTCACCGGAATCAACCTTACCTACTGCTTGG 7374
 QY 7374 AACAGCTGGCATTAAGTCTTTGGCGAGCCCGCCCAAGCGGAGTTTCAAGGCTTCCA 7433
 Db 7375 CCGAGCTTGCACCAAAAGTTTTGGGAGCTCTCAACTCCGCGATTAAGCGGCGACAATA 7434
 QY 7434 CGGGGCGGGCGCTGCGGATTCGCGAGTCCAGACGCTCTCTGATGATGTTGGCCCTTTCGG 7493

Db 7435 TGACAAATCTCTCTG-----AGCCCGCCCTTCTGCTGCCCCCGACTCG 7482
 QY 7494 AGACAGTTTCATCTCTTCCATGCCCCCTCTGAGGGGAGCTTGGAGATCCAGACCTGG 7553
 Db 7483 ACCTTGAAGTCTTATTTCTTCCATGCCCCCTCTGAGGGGAGCTTGGGATCCGGA----- 7537
 QY 7554 AGCCTGAGCAGGTAGAGCCCCCAACCCCGCCCGGGGGGTGGCAGCTCCCGGCTCGG 7613
 Db 7538 -----TTTCAGGAGC 7548
 QY 7614 ACTCGGGGTCTGGTCTACTTGTCTCGAGGAGGAGCTCTCGTCTGTCTCTCATGT 7673
 Db 7549 GGTCTATGTCTGAGCGGTCTAGTAGTGGGGCGGACACGGAAGATCTGCTGTCTCTCAATGT 7608
 QY 7674 CATACTCTGGACCGGGGTCTAATACTCTTGTAGTCTCCGAGAGGAGATTTACCGA 7733
 Db 7609 CTTATACCTGGACAGCGCACTCGTCAACCGTGGCTGCGGAAGAACAAAACTGCCCA 7668
 QY 7734 TTAAACCCCTTGAAGCACTCCCTGTTGCGATATCAAAAGGTGTACTGTACCAACAA 7793
 Db 7669 TCACGCACTGAGCAACTGTTGCTAGCGCATCACAATCTGTATATTTCCACCACTTAC 7728
 QY 7794 AGAGCGCCTCACTAAGGGCTAAAAAGTAACTTTTGTAGGATGCAAGTGTCTCGACTCT 7853
 Db 7729 CGAGTGTCTGCCAAAGGAGGAAAGTCACATTTTACAGACTGCAAGTCTTGGACAGCC 7788
 QY 7854 ACTACGACTCAGTCTTAAAGGACATTAAGCTAGCGGCTTCAAGGTACCGCAAGGCTCC 7913
 Db 7789 ATTACGAGCGTGTCAAGGAGGTCAAGCAGCGCGCTCAAAAGTGAAGGCTTAACCTTGC 7848
 QY 7914 TCACCATCGAGGAGCTTGGCAGTTTAAACCCCACTTCTGCAAGATCTAATATGGT 7973
 Db 7849 TATCGTATAGGAAGCTTGGACCTGACGCCCCCAATTCAGCCAAATTCAGATTTGGCT 7908
 QY 7974 TTGGGGCTAAGGAGTCTCGAGTCTGTCGGGAGGCGCTTAAACCACTCAAGTCCGCTG 8033
 Db 7909 ATGGGGCAAAAGACGCTCTGTCATGCCAGGAGGAGCGCTAGCCCACTCACTCACTCGT 7968
 QY 8034 GGAAGGACTCTTGGAGGACTCAGAAACACCAATTTCCCAACCACTTATGGCAAAATG 8093
 Db 7969 GGAAGAGCTTCTGGAAGACAGTGTAAACCAATAGACACTATCATCTATGCGCAAGAG 8028
 QY 8094 AGGTGCTTCTGCTGAGCACCACCAAGGGGCGCAAGAAAGCAGCTCGCTTATCTGTTAC 8153
 Db 8029 AGGTCTTCTGCTTCAAGCTGGAAGGGGCTGTAGCCAGCTCTCTCATCTGTTTCC 8088
 QY 8154 CTGACCTCGGCTCAGGGTCTGCGAAGAGATGGCCCTTTTATGACATTTACACAAAATTC 8213
 Db 8089 CCGACCTGGCGTGGCGTGTGCGAAGAGATGGCCCTGTACGCTGGTTAGCAAACTCC 8148
 QY 8214 CTCAGCGGTGATGGGGGCTTCTTATGATTCAGTATTTCCCGCTCAGCGGGTAGAT 8273
 Db 8149 CCGTGGCGCTGATGGGAAAGCTCTACGGATTCCTCACTACAGGAGCGGGTTGAAT 8208
 QY 8274 TTTCTTGAAGCATGGGGGAAAGAGGACCTTATGGTTTTTCTGATGATACCCGAT 8333
 Db 8209 TCTCTGCAAGCTGGAAGTCCAAAGAACCCCGATGGGGTCCCGTATGATACCCGT 8268
 QY 8334 GCTTTGACTCAACCGCTCACTGAGAGACATCAGAGCTCAGAGTTCATATATCGGGCT 8393
 Db 8269 GTTTTGACTCCACAGTCACTGAGAGGACATCCGTCGAGGAGGCAATTTACCAATGT 8328
 QY 8394 GCTCTTGGCGAGAGGCGCACCTGCGATACCTCGCTAATCTGAGAGATTTAGTGG 8453
 Db 8329 GTGACCTGGACCCCAAGCGCGTGGCCATCAAGTCCCTCACTGAGAGGCTTTATGTTG 8388
 QY 8454 GAGGCTCTATGTTCAACAGCAAGGCGCAAACTCGGGGTACAGGCGTGGCGCGCAGG 8513
 Db 8389 GGGGCGCTTCTTACCAATTTCAAGGGGGGAAACTGCGGTCTATCGAGGTGCGCGCAGG 8448
 QY 8514 GGGTCTCACCACCTAGCATGGGGAACACCATCATGCTAGCTGAAGACCTTTAGCGGCTT 8573

Db	4135	GTGCTTACATGTCCAAAGCCCATGGGGTGCATCTTAATATCAGACCGGGGTGAGAACAA	4194	5286	TGCCACCTTGCATGCAAGCCGACCTTGAGGTATGACACAGACATGGGTCTTTGGCAGGGG	5345
Qy	4206	TGACAGCCGGGGCCCATCAGCTACTCCACATATGCAATTCCTCGCCGATGGGGCT	4265	5275	TGATGACATGCATGTGCGGCGGACCTGGAGGTGCTCAGAGACCTCTGGTGGTGGG	5334
Db	4195	TTACCACTGGAGCCCATCAGCTACTCCACCTACGCAAGTTCCTTGGCGACGGGGT	4254	5346	GAGTCTTGGCGGCGCTGCGCGGTATGCTGGCGACCGGGTGTGTTGATCATCATCGGCC	5405
Qy	4266	GTGCGGCGCGCTACGACATCATCATATGTGATGAATCCATGCGGTGGACTCTACCA	4325	5335	GGTCTCGCTGCTCTGGCGGGTATTTGCTGTGCAACAGGCTGCGTGGTCTAGTGGGCA	5394
Db	4255	GCTCAGGAGCGCTTATGACATAAATTTGTGACGAGTGCCACTCCACGGATGCCACAT	4314	5406	GCTTGCACATTAACACAGCAGCGCTGTTGCGCGGACAGAGGTCTCTATGAGGCTT	5465
Qy	4326	CCATCTTGGGATCGGAAAGTCTCTGATCAGACAGACAGCTGGGGTCAGACTAACTG	4385	5395	GGATTCCTTGTTCGGGAAAGCCGGCAATTAATCTGACAGGAGGTCTCTACCGAGT	5454
Db	4315	CCATCTTGGGATCGGCACTGTCTTGAACAGCAGAGCTGCGGGGGGAGATTTGGTTG	4374	5466	TTGATGAGATGAGGAAATGTGCTCTAGGGCGGCTCTCATTTGAAGAGGGCGACGGATAG	5525
Qy	4386	TGCTGGCTACAGCTACGCCCTCGGTCAGTGACAAACCCCAACCAATAGAGAGG	4445	5455	TCGATGAGATGGAAGAGTGTCTCTCAGCACTTACCGTACATCGAGCAAGGATGATGCTCG	5514
Db	4375	TGCTCGGCACTGTACCCCTCGGGCTCCGTCACTGTGTCCATCTCTAATCATCGAGAGG	4434	5526	CCGAGATGTGAAGTCCAAAGCTTAAAGGCTTATTCAGAGCAAGCTTCCAAACAGCTCAAG	5585
Qy	4446	TGGCCCTTGGGAGGAGGCGAGATCCCTTCTATGAGGAGGCGATTCCTCTGTTCACA	4505	5515	CTGAGCAGTTCAGCAGAAAGGCCCTCGGCTCTCTGACAGACCGGCTCCGCCATGACAGAGG	5574
Db	4435	TTGCTCTGTCCACCACCGGAGATCCCTTCTACGGCAAGGCTATCCCTCTCGAGTGA	4494	5586	ACATACAAACCCACTGTGCGAGGCTTCAATGCGCCCAAGGTAGAACAAATTCGCGCCAAACACA	5645
Qy	4506	TCAAGGAGGAGACATCTGATCTTCTGCGCATTCAAAGAAAGTGTGACGAGCTCGCG	4565	5575	TTATCAACCCCTGCTGTTCAGACCAACTGGCAGAACTCGAGGTCTTCTGGGCGAGACACA	5634
Db	4495	TCAAGGGGGGAAAGACATCTCATCTTCTGTCACTCAAAAGAAAGTGTGCGAGCTCGCG	4554	5646	TGTGGAACCTTCATTAGCGGCATCCAAATACCTCGCAGGACTATCAACACTGCCAGGGAACC	5705
Qy	4566	CGGCCCTTGGGGTATGGCTTGAACCTCAGTGCGACTACTACAGAGGTTGAGAGTCTCGG	4625	5635	TGTGGAATTTCACTCAGTGGGATACATAATTTGCGGGCTGTCTCAACGCTGCTGTGTAACC	5694
Db	4555	CGAAGCTGTGCGCATTTGGGATCAATCGCTGGCTACTACCGCGACTTTGAGTGTCTG	4614	5706	CTGCACTAGCTTCATGATGCGGCTTCAAGTCCGCGCTCAACAGTCCGCTGTCTCAACAGCA	5765
Qy	4626	TAATACCAACTCAGGAGAGAGTGTGTGTGCTGCGCACCGACCGCTCATGACAGGGTATA	4685	5695	CCGCCATTCCTCATTTGATGGCTTTTACAGCTGCGCTCACAGCCCACTAACCACTGGCC	5754
Db	4615	TCATCCCGACACCGGCGGATTTGTGCTGTGTGTCGACCGATGCTCTCATGACTGGCTTTA	4674	5766	CCACTATCTTCTCAACATTTTGGGGGGTGGCTAGCATTCGCCAAATTTGACACACCCGCGG	5825
Qy	4686	CTGGGACTTTGACTCGCTGATCGACTGCAACAGTACGCGGTCACTCAAGTGTGTAGACTTCA	4745	5755	AAACCTCTCTTCAACATTTGCGGGGGTGGGTGGCTGCCAGCTGCGCGCCCCCGGTG	5814
Db	4675	CCGCGCACTTCGACTCTGTATAGACTGTCAACAGCTGTGTCACTCAGACAGTCSGATTTCA	4734	5826	GGGCCACTGGCTTGTGTGCTAGTGGGCTAGTGGGAGCTGCGGTGAGGAGTATAGGCTTAG	5885
Qy	4746	GTTTAGACCCACATTCACCATACACACAGATTTGCTCAGAGCGCTGTCTCAGCTA	4805	5815	CCGCTACCGCTTTTGGGGGCTGGCTTAGCTAGTGGCGCCGCTATCGGAGCGTGTGGACTGG	5874
Db	4735	GCCTTGACCTTACCTTTACATTTAGAAACACAGCTCTCCCGAGGATGCTGTCTCAGGA	4794	5886	GTAAAGTGTAGTGGACATCTCGGAGGGTATGGTGGGGGCAATTTGGGGGCTCTGCTCG	5945
Qy	4806	GCCAGCCCGGGTGCACGGGTAGGGAAGACTGGGCATTTATPAGTATGTTTTCACCTG	4865	5875	GGAAAGTCTCTGCTGGACATTTCTTGAGGGTATGGCGCGGGCGTGGGGGAGCTCTTGTAG	5934
Db	4795	CTCAGCCCGGGCAGGACTGGCAGGGGAGCCAGGCATCTACAGATTTGTGGCACCGG	4854	5946	CATTCAAGATCATGTCTGCGGAGAACCTTCCATGAGAGATGTGTCACCTCTGCTGCTG	6005
Qy	4866	GTGAGCGAGCTCAGGAATGTTTGAAGTGTAGTGTCTGTGAGTGTCTACAGCGAGGGG	4925	5935	CATTCAAGATCATGAGCGGTGAGGTCCCTTCCAGGAGGACCTGGTCAATCTGCTACCCG	5994
Db	4855	GGGAGCGCCCTCCGGCATGTTGACCTCGTCCGTCTCTGTGAGTGTCTATGACGCGGGCT	4914	6006	GAATTCCTGCTCCCGGTGCTTGGTGTAGTGGGAGTCAATCTGCGCGGCATTTCTGCGCGAC	6065
Qy	4926	CCGATGCTATGAGCTCACACATCGGAGACACCGTCAGGCTCAGGGCGTATTTCAACA	4985	5995	CCATCTCTCGCTCGGAGCCCTTGTAGTGGTGTGCTGCGCAGCAATACTGCGCGCGG	6054
Db	4915	GTGCTTGTATGAGCTCATGCCCCGAGACTACAGTTAGGCTACGAGCGTACATGAAACA	4974	6066	ACGTGGGACCGGGGAGGGCGGCTCAATGAGTGAATAGACTCATTTGCTTGTCTTCCA	6125
Qy	4986	CGCCCGGTTTGCCTGTGTGCAAGACCATCTGTAGTTTGGGAGGCAAGTTTTCACCGGCC	5045	6055	ACGTTGGCGCGGCGAGGGGCGAGTGCATGGAATGAACCGGCTTAATAGCCTTCCGCTCCC	6114
Db	4975	CCCCGGGCTTCCCGTGTGCGAGACCATCTTGAATTTGGGAGGGGCTCTTTTACGGGCC	5034	6126	GAGGAAATCAGCTGCGCCCCCACCCACTAGTGTACGAGTCCGATGCGTGTGCGACGTGTGA	6185
Qy	5046	TCACACATAGATGCCCATCTCTTCCCAACAAAGCAATCGGGGGGAAATTTTCGCAT	5105	6115	GGGGGAACCATGTTTCCCCCAGCACTACGTGCGGAGAGCGATGACGCGCGCGGCTCA	6174
Db	5035	TCACCCATATAGATGSCCACTTTCTATCCAGACAAAGCAGAGTGGGGGAGAACTTTCCTT	5094	6186	CCCAACTACTTGGCTCCCTTACCAATACCGCCTGCTCAGAAAGCTCCCAACTGGATTA	6245
Qy	5106	ACTTAACAGCTTACAGGCTACGTGCGCTAGGGCCAAAGCCCCCCCCCGCTGGG	5165	6175	CTGCCATCTACAGCGCTCACTGTAAACCGCTCTGAGGCGACTACATCAGTGGATAA	6234
Db	5095	ACCTGGTAGCGTACAAGCAACCGGTGCGCTAGGGCTCAGGCGCTTCCCGCATCTGTGGG	5154	6246	CTGAGGACTGCCCATCCCATGCGGGCTCTGCTGCTCGCGATGTGTGGGACTGGGTTT	6305
Qy	5166	ACGTCTATGTGGAAGTGTGACTCGACTCAAGCCCACTGCTGGGGGCCCACTCTCTCC	5225	6235	GCTCGGAGTGTACCACTCCATGCTCCGGCTCTCTGCTTAAGGGACATCTGGGACTGGATAT	6294
Db	5155	ACCAGATGTGGAAGTGTGATCGCTTAAACCCACCTCCATGGGGCCCAACACCCCTGC	5214	6306	GCACCATCTTAACAGAGCTTTAAATTTGGCTGACCTCCCAATTTATTTCCCAAGATGCCG	6365
Qy	5226	TGTACCGTGTGGCTCTGTTACCAACAGAGGTACCCCTCACAACATCTCCGCTGACGAAATACA	5285	6295	GCAGGTGTGAGCGACTTTTAAGACCTTGGCTGAAAGCCAAAGCTCATGCCCAACTGCTCTG	6354
Db	5215	TATACAGACTGGGGCTGTTTCAAGATGAAGTCAACCTGACCGCAACCAATCAACCAATACA	5274			

CC proteins encoded by the genome of hepatitis C virus (HCV), such as NS5A
 CC protein, for replication, and optionally production, of HCV or its
 CC derived viable mutants, in an appropriate culture medium. The cells are
 CC used to produce HCV particles, and to screen for anti-HCV agents.
 CC Inhibitors of prenylation are useful for treating HCV infection. This
 CC polynucleotide sequence represents the DNA encoding an HCV protein
 CC relating to the invention.
 XX

SQ Sequence 9622 BP; 1889 A; 2897 C; 2719 G; 2117 T; 0 other;

Query Match 46.8%; Score 4543.6; DB 25; Length 9622;
 Best Local Similarity 67.4%; Pred. No. 0;
 Matches 6550; Conservative 0; Mismatches 3054; Indels 114; Gaps 6;

Qy	6	CCCTAATAGGGGACACTCCGCATGATGATGAGTGTCTGCTGAGGAACTTCTTCTTCA	65
Db	7	CCCTGATGGGGGACACTCCACATAGATGATGAGTGTCTGCTGAGGAACTTCTTCTTCA	66
Qy	66	CGCAGAAAGCGTCTAGCCATGGGCTTAGTATGAGTGTCTGCTGAGGAACTTCTTCTTCA	125
Db	67	CGCAGAAAGCGTCTAGCCATGGGCTTAGTATGAGTGTCTGCTGAGGAACTTCTTCTTCA	126
Qy	126	TCCCGGAGAGCATAGTGTCTGCGAAACCGGTGAGTACACCGGAATTCGCGGAGAC	185
Db	127	TCCCGGAGAGCATAGTGTCTGCGAAACCGGTGAGTACACCGGAATTCGCGGAGAC	186
Qy	186	TGGGTCTCTTCTGGATTAACCCACTCTATGCGCGGCAATTCGGGCTGCCCGCAAGA	245
Db	187	CGGGTCTCTTCTGGATTAACCCCGCTCAATGCTTGGAGATTTGGGCTGCCCGCAAGA	246
Qy	246	CTGCTAGCCGAGTAGTGTCTGGGTTGCGAAAGGCTTGTGCTGCTGATGAGGCTCTT	305
Db	247	CTGCTAGCCGAGTAGTGTCTGGGTTGCGAAAGGCTTGTGCTGCTGATGAGGCTCTT	306
Qy	306	GGGAGTCCCGGGAGTCTCGTAGACCGTGCACCATGAGCACAAATCTTAAACCTCAA	365
Db	307	GGGAGTCCCGGGAGTCTCGTAGACCGTGCACCATGAGCACAAATCTTAAACCTCAA	366
Qy	366	GAAACCAAAAGAAACACCAACCGTGCAGCAAGAGCTTAAATTTCCGGGCGGGCC	425
Db	367	GAAACCAAAAGAAACACCAACCGTGCAGCAAGAGCTTAAATTTCCGGGCGGGCC	426
Qy	426	AGATCGTTGGCGGAGTACTTGTTCGCGGAGGCGCCAGGTTGGGTTGGCGCGA	485
Db	427	AGATCGTTGGCGGAGTACTTGTTCGCGGAGGCGCCAGGTTGGGTTGGCGCGA	486
Qy	486	CAAGGAAGACTTCGGAGCGTCCAGGACAGTGGAGCGCGACCCATCCCTAAAGATC	545
Db	487	CAAGGAAGACTTCGGAGCGTCCAGGACAGTGGAGCGCGACCCATCCCTAAAGATC	546
Qy	546	GGGCTCCACTGGCAATCTGGGGAACACAGGATACCCCTGCGGCGCTATACGGGAATG	605
Db	547	GTGGGCTCCAGGAGGAGGAGCTTGGGCTCAGCCCGGGTACCCCTTATGGCAATG	606
Qy	606	AGGAGCTCGGCTGGGAGGAGGCTTCTGTCCTCCCGGAGGTTCCCGCTCTTGGGGCC	665
Db	607	AGGAGCTCGGCTGGGAGGAGGCTTCTGTCCTCCCGGAGGTTCCCGCTCTTGGGGCC	666
Qy	666	CCAAATGACCCCGGATAGTGTGCGCAACGTTGGGTAAGGTCATGATACCCCTAAAGTGC	725
Db	667	CCAAATGACCCCGGATAGTGTGCGCAACGTTGGGTAAGGTCATGATACCCCTAAAGTGC	726
Qy	726	GCTTTGCGGACCTATGGGTAACATCTGTCGTTGGGCGCCCGCTCGGGGCGTCCCA	785
Db	727	GCTTTGCGGACCTATGGGTAACATCTGTCGTTGGGCGCCCGCTCGGGGCGTCCCA	786
Qy	786	GAGCTCTCGCATGGGCTGAGAGTCTCTGGAGGACGGGGTTAAATTTGCAACAGGGAAT	845
Db	787	GGGCGCTGGCGCATGGGCTCGGGTTCTGGAAGACGGCGTGAATCTATGCAACAGGGAAT	846
Qy	846	TACCGGTTGCTCTTTTCTATCTTCTGTCGTTGGGCTCTGCTCTGATACACACCGG	905
Db	847	TTCCTGGTTGCTCTTTCTATCTTCTTCTGTTGGGCTCTGCTCTGCTGACTGTGCGG	906

Qy	906	TCTCGGCTGCCGAAGTGAAGAACATCAGTACCGGCTACATGTTGTTAAACGACTGCACCA	965
Db	907	CTTACGCTTACCAAGTGGCGCAATCTCGGGCTTTTACCATGTCACCAATGATGCGCTA	966
Qy	966	ATGACAGCAATTAACCTGGAGCTTCAGGCTGTCTCTCAGCTCCCGGGTGGTCCCGT	1025
Db	967	ATTGAGTATTGTGTACGAGGGCGGATGCCATCTCTGCACACTCCGGGTTGTCTCTT	1026
Qy	1026	GGGAGAACTGGGGAATCATCTCAGTGTGTGATACCGGCTCTCACGAATGTGGCGCTGC	1085
Db	1027	GGTTTCGAGAGGTAAAGCTCTGAGTGTGTGGTGGCGGTGACCCCAAGTGGCCACCA	1086
Qy	1086	AGCGGCGGCGGCTCTCACGAGGCTTCGCGACGCAATGCAATGTTGTGATGTCG	1145
Db	1087	GGGAGGCAAACTCCCCACAAACGAGCTTCGAGCTCATATCGATCTGCTTGTGGGAGCG	1146
Qy	1146	CAACGCTCTGCTGTCGCTCTACGTTGGGGAACCTCTGCGGTGGGGTGTCTCGACGCC	1205
Db	1147	CAACGCTCTGCTGTCGCTCTACGTTGGGGAACCTCTGCGGTGGGGTGTCTGTTTGTGGTC	1206
Qy	1206	AAATGTTCAATGTTCTCGCGGAGCACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	1265
Db	1207	AAATGTTCAATGTTCTCTCCAGGCGCACTGGAAGCGCAAGCTGCAATTTGTTCTTCT	1266
Qy	1266	ACCTGTTACCATCACTGGACACCGCATGGCATGGGATGATGATGATGATGATGATGATGAT	1325
Db	1267	ATCCCGGCGCATATAACGGGTCATCGCATGGCATGGGATGATGATGATGATGATGATGAT	1326
Qy	1326	CGGCTACCATGATCTTGGGCTAGCGCATGGGCTGTCGCGGAGTCAATATAGACATATTA	1385
Db	1327	CGGCAAGCTTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1386
Qy	1386	GGGCGCTCATTTGGGCGCTCATGTTGCGGTGTCCTTCTATGTCAGAGGAGCGTGGG	1445
Db	1387	CTGTGCTCTCTGGGAGTCTCTGGGCGCATAGGCTATTTCTCATGTTGGGAACTGGG	1446
Qy	1446	CGAAAGTCTGTGTCATCTTCTGTTGGCGCGCGGTTGAGCGCGCAACCATCATCTGTTG	1505
Db	1447	CGAAAGTCTGTGTCATCTTCTGTTGGCGCGCGGTTGAGCGCGCAACCATCATCTGTTG	1506
Qy	1506	GGGTTCTGCGGCGAGACCGCGGCTCACAGCTTATTTGATGATGAGGCGGCGGCGG	1565
Db	1507	GGGAAAGTGGCGGCGACACCGGCTTGGTCTCTTCTTACACGAGGCGGCGGCGG	1566
Qy	1566	AGAAATCCAGCTCGTTTAAACCAATGCGAGCTGGGACATCAACCGCACCGGCTTGAAT	1625
Db	1567	AGAAATCCAGCTCGTTTAAACCAATGCGAGCTGGGACATCAACCGCACCGGCTTGAAT	1626
Qy	1626	GCAATGACTCTTGGCACACCGGCTTATCGCGTCTCTGTTTCTACACCCAGGCTTCAACT	1685
Db	1627	GCAATGACTCTTGGCACACCGGCTTATCGCGTCTCTGTTTCTACACCCAGGCTTCAACT	1686
Qy	1686	CGTCAGATGTCGCGAACGATGTCGCTGCGCGAGTATCGAGGCTTCCGGTGGGAT	1745
Db	1687	CTTCAGGCTGCTCGAGAGGTTGGCGAGCTTGGCGAGCTTACCGATTTTTCGCCAGGGCT	1746
Qy	1746	GGGCGCTTGGCAATATGAGGATAATGTCAACCAATCCAGAGGATATGAGACCTTATGCT	1805
Db	1747	GGGCTCCATCATGTTATGCCAACGGAAGCGGCTT-----GACGAACGCGCTTACTGTT	1800
Qy	1806	GGCACTACCCCAAGGCGAGTGTGGGCTGCTCCGCAAGAGCTGTGTGTGGGCGGAGTGT	1865
Db	1801	GGCACTACCCCTCAAGACCTTGTGGCATTTGTGCCCGCAAGAGGCTGTGTGGGCGGAT	1860
Qy	1866	ACTGTTTACCCCGGAGGCTGTGGGCGAGCGGAGCTTGGAGCGGCGGCTT	1925
Db	1861	ATTGCTTACTCCCGGCGGCTGTGGTGGGAAAGGAGCGGAGCTGTGGGCGGCTTACT	1920
Qy	1926	ACAGTGGGCGGAGATGAGACAGATGCTCTTCTTATGAAACAGACATCGACACCGCTGG	1985
Db	1921	ACAGTGGGCGGAGATGAGACAGATGCTCTTCTTATGAAACAGACATCGACACCGCTGG	1980

Db 8329 GTACCTGGACCCCAAGCCCGGTGGCCATCAAGTCCCTCACTGAGAGCTTTATGTTG 8388
 Qy 8454 GAGGGCTATGTTCAACAGCAAGGGCCAAACCTGCGGTACAGGCGTTCGCCGCCAGCG 8513
 Db 8389 GGGGCCCTCTTACCAATTCAGGGGGGAAACCTGCGGTACCGAGGTGCCGCCAGCG 8448
 Qy 8514 GGGTGTCTCACTAGCATGGGGAACACCATCATCATGCTACGTAAAGCCCTTAGCGGCTT 8573
 Db 8449 GCGTACTGACAACTAGCTGTGGTAAACCCCTCACTTGTCTACATCAAGGCCGCCGAGCTT 8508
 Qy 8574 GTAAGCTGAGGATAATCGGCCCAACATGCTGTATGCGCGCATGACTTGTGTTGCA 8633
 Db 8509 GTCGAGCCGAGGGCTCCAGGACTGACCATGCTCGTGTGTGGCGACGACTTAGTCTGTTA 8568
 Qy 8634 TCTCAGAAAGCCAGGGGACCGAGGAGGAGCGGAACCTGAGAGCTTTCACGGAGCTTA 8693
 Db 8569 TCTGTGAAGATGCGGGGCTCCAGGAGGAGCGCGGAGCTTTCACGGAGGCTTA 8628
 Qy 8694 TGACCAAGTATCTGCCCCCTCTCTGTTGACCCCCCAGAGCCGGAGTATGATCTGGAGCTGA 8753
 Db 8629 TGACCAAGTATCTGCCCCCCTCGGGGACCCCCCACAACAGATACGACTTGGAGCTTA 8688
 Qy 8754 TAACATCTTGTCTCTCAAAATGTTGTGTGGCGTGGGCCCAACAGGCCGCCGAGATCT 8813
 Db 8689 TAACATCATGCTCTCTCAACGTTGTAGTGTGCCCAACGACGCGCTGGAAGAGGGTCTACT 8748
 Qy 8814 ACCTGACCAAGACCTTACCATCTCAATGCGCCGCTGCTGGGAACAGTTAGACACT 8873
 Db 8749 ACCTTACCCGCTGACCTCAACACCCCTCGGAGAGCCGCTGGGAGACGACGAACACA 8808
 Qy 8874 CCCTGTCAATTCATGGCTGGGAACATCATCAGTACGCCGCCGACCATATGGGCTCGCA 8933
 Db 8809 CTCAGTCAATCTCTGGCTAGGCAATATATCATGTTTGGCCCCCACTGTGGCGAGGA 8868
 Qy 8934 TGGTCTCTGATGACACACTTCTTCTCATCTCATGCTCAAGACACGCTGACCAAGCC 8993
 Db 8869 TGATACATGATGACCACTTCTTTAGCGTCTCATAGCCAGGATCAGCTTGAACAGGCTC 8928
 Qy 8994 TCNACTTTGAGATGATGAGGAGCGGTCTACTCGGTAGTCCCTTGGACCTCCAGCTATAA 9053
 Db 8929 TTAACCTGTGAGATCTACGAGCGCTGTCTTCCATAGAACCACTGGATCTACCTCCAATCA 8988
 Qy 9054 TTGAAGGTTTACATGGCTTACGCTTTCTCTGCACACATACATCTCCCAAGCACTCA 9113
 Db 8989 TTCAAGAGCTCCATGSCCTCAGCGCATTTTCACTCACAGTTACTCTCAGGTGAATCA 9048
 Qy 9114 CAGGGTGGCTTCAAGCCCTCAGAAAATCTGGGGCCCAACCCCTCAGAGCGTGAAGAGCC 9173
 Db 9049 ATAGGGTGGCCGATGCTCAGAAAATCTGGGGTCCCGCTTGGCGAGCTTGGAGACAC 9108
 Qy 9174 GGGCAGTGCAGTCAAGGCGTCCCTCATCTCCGTTGGGGGAGAGCGCGTTTGGGTC 9233
 Db 9109 GGGCCCGAGCGTCCCGCTTAGGCTTCTGTCCAGAGAGGAGGCGGTGCTCATATGCGCA 9168
 Qy 9234 GATATCTCTTCAATTTGGGGGTTGAGAACCAAGCTCAAACTCACTCCATTTGCCGGAAGCGC 9293
 Db 9169 AGTACTCTTCACTGGGAGTGAAGAAAGCTCAAACTCACTCAATAGCGCGCTG 9228
 Qy 9294 GCTCTCTGATTTATCCAGTGTTCACCGTCCGCGCCGCGGGCGGAGCATTTATCACA 9353
 Db 9229 GCGGCTGAGTCTGTCGGTGTGTTTCAAGGCTTACAGCGTGGCTTACAGCGGGGAGCAITTTATCACA 9288
 Qy 9354 GCGTGTGCGTGGCCGACCCCGTTATGCTCTTTGGCTTACTCTTCTTTTGTAGGGG 9413
 Db 9289 GCGTGTCTCATGCGCGCCCGCTGTTTCTGGTTTTCGCTTACTCTCTGCTCGCTGAGGGG 9348
 Qy 9414 TAGGCTCTTCTTACTCCCGCTCG--GTAGAGCGGCACACATTAAGTACTCTCATATG 9471
 Db 9349 TAGGCACTACTCTCTCCCAACCGATGAAGTTGGGTAAACACTCCGCGCTCTTAGGC 9408
 Qy 9472 TAAGTGTCCCTTT 9531
 Db 9409 CATTTCTGTT 9468

Qy 9532 TT 9591
 Db 9469 TTTCTTT 9528
 Qy 9592 ACCTTATTTTACTTTCTTTCTGCTGGTCCATCTTAGCCCTAGTCAACGCTAGCTGTGA 9651
 Db 9529 TTTCTTTTCTTCTTTCTTTTAAATGGTGGCTCCATCTTAGCCCTAGTCAACGCTAGCTGTGA 9588
 Qy 9652 AAGTTCCTGAGCCGAGTGTGAGAGTGCAGAGTGCCTTAATGCTCTCTGAGATCATGT 9711
 Db 9589 AAGTTCCTGAGCCGATGACTGACAGAGTCTGATCTGCTCTCTGAGATCATGT 9648
 RESULT 15
 AAL54424
 ID AAL54424 standard; DNA; 9622 BP.
 XX
 AC AAL54424;
 DT 03-APR-2003 (first entry)
 XX Hepatitis C virus DNA, SEQ ID No 2.
 XX Virucide; hepatotropic; antiinflammatory; prenylation; hepatitis C virus;
 KW HCV; NS5A protein; anti-HCV agent; HCV infection; gene; ds.
 XX Hepatitis C virus.
 FH Key Location/Qualifiers
 CDS 342..9377
 FT /*tag= a
 FT /product= "HCV protein"
 FT /transl_except= (pos:3843..3845, aa:His)
 FT /transl_except= (pos:3846..3848, aa:Ala)
 FT /transl_except= (pos:3849..3851, aa:Val)
 FT /transl_except= (pos:3852..3854, aa:Gly)
 FT /transl_except= (pos:3855..3857, aa:Leu)
 FT /transl_except= (pos:3858..3860, aa:Phe)
 FT /transl_except= (pos:3861..3863, aa:Arg)
 FT /transl_except= (pos:3864..3866, aa:Ala)
 FT /transl_except= (pos:3873..3875, aa:Cys)
 FT /transl_except= (pos:3876..3878, aa:Thr)
 FT /transl_except= (pos:3879..3881, aa:Arg)
 FT /transl_except= (pos:3882..3884, aa:Gly)
 FT /transl_except= (pos:3885..3887, aa:Pro)
 FT /transl_except= (pos:3888..3890, aa:Leu)
 FT /transl_except= (pos:3891..3893, aa:Leu)
 FT /transl_except= (pos:3897..3899, aa:Pro)
 FT /transl_except= (pos:3900..3902, aa:Ala)
 XX
 PN FR2824072-AL.
 XX
 PD 31-OCT-2002.
 XX
 PF 27-APR-2001; 2001PR-0005732.
 XX
 PR 27-APR-2001; 2001PR-0005732.
 XX
 PA (CNRS) CNRS CENT NAT RECH SCI.
 XX
 PI Wychowski C, Duverlie G, Dubuisson J, Pillez A;
 XX WPI; 2003-142456/14.
 DR P-PSDB; AAO26784.
 XX
 PT Using cells that can prenylate proteins for replication and production
 of hepatitis C virus (HCV), useful for screening compounds for anti-HCV
 activity -
 PS Claim 10; Fig 2; 85pp; French.
 CC The invention relates to cells which are able to cause prenylation of

[illegible]

Db 4015 TGGCCCACTGCGATGCTCCCAACGGGAGCGGTAAAGACACCAAGGTCCTGGCTGCGTAGC 4074
Qy 4086 CTGCTCAGGGGTATAAAGTGTAGTCTTAATCCCTCAGTGGGTGCGACCCCTGGGGTTTG 4145
Db 4075 CAGCCAGGCTACAGGTGTGGTGTCAACCCCTCTGTGTCTGCAACGCTGGGCTTTG 4134
Qy 4146 GGGGTAATTTGTTAAGGCAATGGGCAATCAATCCCAACATTAGGACTGGAGTCAGACTG 4205
Db 4135 GTGCTTACATGTCCAAAGGCCATGGGTTGATCCTAATATCAGGACCGGGGTGAGAACA 4194
Qy 4206 TGAGGACCGGGGCGCCCATCAGTACTCCACATATGCGCAATTCCTCGCGATGGGGCT 4265
Db 4195 TTACCACTGGCAACCCCATCAGTACTCCACCTACCGCAAGTTCTTGGCGACGGCGGT 4254
Qy 4266 GTGCGGCGCGCTACGACATCATATATGATGAATGCCATGCCGTGGACTCTACCA 4325
Db 4255 GCTCAGGAGTGTATGACATATATATTTGTGACGAGTGCCACTCAACGATGCCACAT 4314
Qy 4326 CCATCTTGGCATCGGAACAGTCTTGTATCAAGCAGAGACGCTGGGGTCAGACTAACTG 4385
Db 4315 CCACTTGGGCATCGGCACCTGTCTTGACCAAGCAGAGACTGCGGGGCGAGACTGGTTG 4374
Qy 4386 TGCTGGCTACAGTACGCCCCCTGGGTGAGTGAACAACCCCAACCAACATAGAGAGG 4445
Db 4375 TGCTGGCACTGTACTCCCTCGGGCTCCGTCACTGTGTCCCATCCATAACATCGAGAGG 4434
Qy 4446 TGGCCCTTGGGAGGAGGCGAGATCCCTCTTATGGAGGGGATTCCTCTGTCTTACA 4505
Db 4435 TTGCTTGTCCACACCGGAGAGATCCCTTTTACCGAAGGTATCCCTCTCGAGGTGA 4494
Qy 4506 TCAAGGAGGAGACATCTGATCTTCTGCAATTCAAAGAAAGTGTGACGACTCGCG 4565
Db 4495 TCAAGGGGAGAGACATCTCATCTCTGCGCACTCAAGAGAGTGGCAGACTCGCG 4554
Qy 4566 CGGCCCTTGGGATAGGGCTTGAATCAGTGGCATCTACAGAGGTTGGAGTCTCG 4625
Db 4555 CGAAGCTGTGCGATTTGGGCATCAATGCCGTGGCTACTACCGCGTCTTGAAGTCTG 4614
Qy 4626 TAATACCACTCAGGAGAGTGTGTGTGTCGCCACCGACGCTCATGACAGGATATA 4685
Db 4615 TCATCCGACACGGGGATTTGTGTGTGTGTGACCGATGTCTCATGACTGGCTTTA 4674
Qy 4686 CTGGGACTTTGATCCGTGATCAGTGCACGATAGCGGTCACTCAAGTTGTAGACTTCA 4745
Db 4675 CCGGCACTTCGACTCTGTATAGACTGCAACACGCTGTCTACAGACAGTCGATTTCA 4734
Qy 4746 GTTTAGACCCCACTTCAACATTAACCAACAGATTTGTCCCTCAAGAGCTGTCTACGTA 4805
Db 4735 GCCTTGACCCCTTACCTTTACATTTGAGACCAACACGCTCCCGGAGTGTCTCCAGGA 4794
Qy 4806 GCCAGCGCGGGTTCGACGGTAGGGAGAGCTGGGCATTTATAGGTATGTTCCACTG 4865
Db 4795 CTCACGCCGGGAGAGACTGGCAGGGGAGCCAGGCATCTACAGATTTGTGGCACCG 4854
Qy 4866 GTGAGGAGGCTCAGGAATTTTGACAGTGTAGTGTCTGTGAGTGTCTACGACGAGGG 4925
Db 4855 GGGAGCGCCCTCCGGCATGTTGCACTCTGTCGTCTCTGTGAGTGTATGACGCGGCT 4914
Qy 4926 CCGCATGTATGAGTCAACCATCGAGACACCGTCAAGGCTCAGGGGTATTTCAACA 4985
Db 4915 GTGCTTGGTATGAGTCAACCGCCCGGAGACTACAGTTAGGCTTACGAGGCTACATGAACA 4974
Qy 4986 CGCCCGGTTTGGCTGTGTCAGACCATCTTGTGAGTTTGGAGGCGAGTTTTCACCGGC 5045
Db 4975 CCCCGGGCTTCCGTGTGCGAGACCATCTTGAATTTTGGGAGGGCGTCTTACGGGCC 5034
Qy 5046 TCACACATPAGATGCCCATCTCTTTTCCCAACAAAGCAATCGGGGAAATTTTCGCAT 5105
Db 5035 TCACTCATATAGATGCCACTTTCTATCCAGACAAAGCAGAGTGGGGAGAACTTTCTCT 5094
Qy 5106 ACTTAACAGCTTACAGGCTACAGTGTGCGTATAGGGCCAAAGCCCCCGCTCGTGG 5165
Db 5095 ACCTGTAGCGTACCAAGCCCGTGTGGCTAGGGCTCAAGCCCCCTCCCGCATCTGTGG 5154

Qy 5166 ACGTCATGTGGAAGTGTGTAAGTCTCAAGCCACACTCGTGGGCCCAACACCTCTCC 5225
Db 5155 ACCAGATGTGGAGTGTGTGATCCGCTTAAACCCACCCCTCCATGGGCCAAACCCCTGC 5214
Qy 5226 TGTACCGCTTGGGCTCTGTGTACACAGAGGTCAACCTTCAACATCCCGTGTAGCAATACA 5285
Db 5215 TATACAGACTGGGCGCTGTTCAGAAATGAAGTCAACCTGACGCAACCCAAATCACCAATACA 5274
Qy 5286 TCGCCACCTGCATGCAAGCGGACCTTGAAGTCAATGACAGCAGCATGGGTCTTGGCAGGG 5345
Db 5275 TATGACATGCAATGTGCGCCGACCTGGAGGTCTGTCAGAGCACCTGGGTGCTGTGGGG 5334
Qy 5346 GAGTCTTGGCGCGCTGCGCGCTATTTGCTCGCACACCGGTGTGTTTCATCATCGGCC 5405
Db 5335 GGCTCTGGCTGCTGTGGCGGCTATTTGCTGTCAACAGGCTGCGGTGTCATAGTGGCA 5394
Qy 5406 GCTTGCACATTAACAGAGAGCGCTGTTGCGCCGAGCAAGAGGTCTCTATAGAGCTT 5465
Db 5395 GGAATGTCTGTCCGGAAGCGGCAATTTATACCTGACAGGAGGTCTCTTACAGGAGT 5454
Qy 5466 TTGATGAGATGAGGAAATGTGCTCTAGGGCGGCTCTATTGAAGGGGCGAGCGATAG 5525
Db 5455 TCGATGAGATGGAAGAGTCTCTCAGCACTTACCGTACATCGAGCAAGGATGATGCTCG 5514
Qy 5526 CCGAGATGCTGAAGTCCAAGATCCAAGGCTTATTCAGCAAGCTTCCAAACAGGCTCAAG 5585
Db 5515 CTGAGCAGTTCAGCAGAGGCGCTCGGCTCTCTGAGACCCGCTCCCGCAAGCAGAG 5574
Qy 5586 ACATACAAACCCACTGTGAGGCTTATGSCCCCAAGGTAGAAACAATTTCTGGGCCAAACA 5645
Db 5575 TTATCAACCCCTGCTGTCCAGACCAACTGGCAGAAATCTGAGGTCTTCTGGCGAAGACA 5634
Qy 5646 TGTGGAATCTTATAGCGCATCCATATCTCGAGGACTATCAACACTGCCAGGAAACC 5705
Db 5635 TGTGGAATTTTCACTAGTGGGATACAATATCTTGGCGGCGCTGTCAACGCTGCTGGTAAAC 5694
Qy 5706 CTGCACTAGCTTCAATGATGGGCTTCACTGCGCCCTCAACAGTCCGCTGTCAACAGCA 5765
Db 5695 CCGCCATGCTTCAATGATGGCTTTTACAGTCCGCTCACCAGCCCACTAACCACTGGCC 5754
Qy 5766 CCACTATCCTTCTCAACATTTTGGGGGCTGTAGCTAGCTCCCAATTCACCAACCGCGG 5825
Db 5755 AAAACCTCTCTTCAACATATTTGGGGGCTGGTGGTCCAGCTCGCGCCCGCGGTG 5814
Qy 5826 GGGCCACTGGCTTCTGTGTAGTGGGCTGTGGAGTGTGCGTGGCAGTATAGGCTTAG 5885
Db 5815 CCGCTACCGCTTGTGGGCGCTGCTTAGCTGGCGCGCCCATCGGACGCTTGGACTGG 5874
Qy 5886 GTAAAGTGTAGTGACATCTTGGCAGGATGTGTGGCGGCTTTCGGGGCTCTCTGTCG 5945
Db 5875 GGAAGTCTCTGTGACATCTTGGAGGATGTGGCGGGGCTGGGGAGTCTTGTAG 5934
Qy 5946 CATTCAGATCATGTCTGGCAGAGCCCTTCAATGGAGATGTGCTCAACTTGTGCTG 6005
Db 5935 CCTTCAAGATCATGAGCGGTGAGGTCCCTCCCAACGAGGACCTGCTCAATCTGTGCGCG 5994
Qy 6006 GAAATCTGTCTCGGGTGTGTGTAGTGGAGTCACTGCGCGGCTCATCTGCGCGGCTAG 6065
Db 5995 CCACTCTCTCGGCTGGAGCCCTTGTAGTGTGCTGTGGCGCAGCAATPACTGCGCGCGC 6054
Qy 6066 AGTGGGACCGGGGAGCGCGCTTCAATGATGAATAGACTCATTTGCTTTTGTCTTCCA 6125
Db 6055 AGTGTGGCGCGGCGAGGGGCGAGTGAATGATGAATGAATGCGGTATATAGCTTCTGCTCCC 6114
Qy 6126 GAGGAAATCACTGCGCCCGCAACCCACTAGTGAACGAGTTCGATGCTGCGAGCGGTGA 6185
Db 6115 GGGGAAACCATGTTTCCCGCAACGCACTAGTCCGCGAGAGCGATGCGAGCGCGCGCTCA 6174
Qy 6186 CCCAACTACTTGGCTTCCCTTACCAACAGCGCTGCTCAGAGACTCCACACTCGATTA 6245
Db 6175 CTGCATACTCAGCAGCTCACTGTAAACCCAGCTCTCTGAGGCGACTGATCATAGTGATTA 6234

[illegible]

CC p30/HCVflong pU, harboring full-length hepatitis C virus (HCV)
 CC cDNA which can be transcribed to produce infectious HCV RNA
 CC transcripts. It is deposited as ATCC 97879. The infectious RNA
 CC can be used in claimed methods for: identifying a cell line that
 CC is permissive for infection with HCV; infecting an animal with HCV;
 CC and transducing an animal susceptible to HCV infection with a
 CC heterologous gene. The invention relates to the determination of
 CC functional HCV genomic RNA sequences, to construction of infectious
 CC HCV DNA clones (see AAV59361), and to the use of the clones, or their
 CC derivatives, in therapeutic, vaccine and diagnostic applications.
 CC The invention is also directed to HCV vectors, e.g. for gene
 CC therapy or gene vaccines. The products and methods can also be
 CC used for studying HCV infection, isolating functional components of
 CC HCV, and for screening for agents capable of modulating HCV
 CC replication in vitro and in vivo.

XX
 SQ Sequence 12980 BP; 2593 A; 3859 C; 3633 G; 2895 T; 0 other;

Query Match 47.6%; Score 4620; DB 19; Length 12980;
 Best Local Similarity 67.8%; Pred. No. 0;
 Matches 6593; Conservative 0; Mismatches 3035; Indels 92; Gaps 6;

QY 6 CCCTTAATAGGGGCGACACTCCGCATGAATCACTCCCTGTGAGGAATCTACTGTCTTCA 65
 DB 7 CCCTGTATGGGGCGACACTCCACCATGAATCACTCCCTGTGAGGAATCTACTGTCTTCA 66

QY 66 CGCAGAAAGCGTCTAGCCATGGCGTTAGTATGATGTCGTACAGCTCCAGGCCCGCCCC 125
 DB 67 CGCAGAAAGCGTCTAGCCATGGCGTTAGTATGATGTCGTGACAGCTCCAGGACCGCCCC 126

QY 126 TCCCGGAGAGCCATAGTGTCTGCGGAACCGGTGAGTACACCGGAATTCGCGGAAGAC 185
 DB 127 TCCCGGAGAGCCATAGTGTCTGCGGAACCGGTGAGTACACCGGAATTCGCGGAAGAC 186

QY 186 TGGGTCTTTCTTTGGATAAACCCACTCTATGCCCGGCCCAATTTGGGGGTGCGCGCGA 245
 DB 187 CGGGTCTTTCTTTGGATAAACCCCGCTCAATGCTGCGAGATTTGGGGGTGCGCGCGA 246

QY 246 CTGCTAGCCAGTAGTGTGGTGTGGAAGGCGTGTGTTGTTGTTGTTGTTGTTGTTGTT 305
 DB 247 CTGCTAGCCAGTAGTGTGGTGTGGAAGGCGTGTGTTGTTGTTGTTGTTGTTGTTGTT 306

QY 306 GCGAGTGTCCCGGGAGTCTCTGAGACCGGTGACCATGAGCACAATCTTAAACCTCAAA 365
 DB 307 GCGAGTGTCCCGGGAGTCTCTGAGACCGGTGACCATGAGCACAATCTTAAACCTCAAA 366

QY 366 GAAAAACCAAAAGAAACCAACCGTGGCGCAACAGAGTTAAGTTTCCGGGGCGCGGC 425
 DB 367 GAAAAACCAAAAGAAACCAACCGTGGCGCAACAGAGTTTCCGGGGCGCGGC 426

QY 426 AGATCGTGGCGAGTATCTGTTGCGCGCAGGGGCGCCAGGTTGGGTGTCGCGCGA 485
 DB 427 AGATCGTGGCGAGTATCTGTTGCGCGCAGGGGCGCCAGGTTGGGTGTCGCGCGA 486

QY 486 CAAGGAAGACTTCGAGCGGTCCAGCAGCTGGAAGGGCGCCAGCCATCCCTTAAAGATC 545
 DB 487 CGAGGAAGACTTCGAGCGGTCCAGCAGCTGGAAGGGCGCCAGCCATCCCTTAAAGATC 546

QY 546 GGGCGTCCACTGCAAACTCTGGGGAACCAAGATACCCCTGGCGCCCTTATACGGAATG 605
 DB 547 GTCCGCGCAGGAGCAGGACCTGGGCTCAGCCCGGGTACCTTGGCGCCCTTATGGCAATG 606

QY 606 AGGGACTCGGCTGGGAGGATGCTCTGTTCCCGCAGGGTCCCGTCCCTCTTGGGGCC 665
 DB 607 AGGGTTGCGGGTGGGAGGATGCTCTGTTCCCGCAGGGTCCCGTCCCTCTTGGGGCC 666

QY 666 CCAATGACCCCGGCACTAGTTCGCGCAACGTGGGTAAAGGTATCATGATACCTTAACTGG 725
 DB 667 CCACAGACCCCGGCGTAGTTCGCGCAATTTGGGTAAAGGTATCATGATACCTTAACTGG 726

QY 726 GCTTTGCGGACCTCATGGGTATACCTCTGTGTTGGGCGCGCCCGCTCGCGGCGTCCCA 785
 DB 727 GCTTTGCGGACCTCATGGGTATACCTCTGTGTTGGGCGCGCCCGCTCTTGGAGGCGCTGCCA 786

QY 786 GAGCTCTCGCGCATCGCGTGAAGTCTTGGAGACGGGTAAATTTTGCAACAGGAACT 845
 DB 787 GGGCCCTGGCGATGGCGTCCGGTTTCGGAAGACGGGTGAATATGCAACAGGAACT 846

QY 846 TACCCGGTTGTCTCTTTCTATCTTCTTGTGGCCCTGTCTCTCTGCAATCACCACCGCG 905
 DB 847 TTCTCGTTGTCTCTTTCTATCTTCTTGTGGCCCTGTCTCTCTTGTGGCGTGGCG 906

QY 906 TCTCGCTGCGAAGTGAAGAACATAGTACCGGCTACATGTGTAGTATACGATGACCA 965
 DB 907 CTTACGCTTACCAAGTGGCAATCTCTCGGGCTTTTACCATGTACCAATGATGTCCTA 966

QY 966 ATGACAGCATTAACCTGGCAGCTCCAGGCTGTCTCCAGCTGCCCGGCTGCTCCGCT 1025
 DB 967 ACTCAGTATGTGTACGAGGCGCGCATCTGACACTCGGGGTGTGCTCTT 1026

QY 1026 GCGAAGAAAGTGGGAATCATCTCAGTGTGGATACCGGTCTTCACGGAATGTGGCGTGC 1085
 DB 1027 GCGTTTCGGAGGGTAAACGCTCGAGGTGTGGGTGGGTGACCCCGCACGCGTGGCA 1086

QY 1086 AGCGGCCCGCGCCCTCAGCAGGCGTTCGGAACGACATCGACATGTTGTGTGATGTCG 1145
 DB 1087 GGGACGGCAAACTCTCCCAACACGAGCTTCGACGTCATATCGATCTGCTTGTGCGAG 1146

QY 1146 CCACGCTCTGCTCTGCCCTCTACGTGGGGACCTCTGCGGTGGGTGATGCTCGAGCC 1205
 DB 1147 CCACCTCTGCTCGGCCCTCTACGTGGGGACCTGCTGGGGTCTGCTCTTCTTGTGTC 1206

QY 1206 AAATGTTCAATGTCTCGCGCAGCAGCACTGTTTTGTCCAAGACTGCAATGCTCATCT 1265
 DB 1207 AACTGTTTACCTTCTCTCCAGCGCCTTCGACGACGACGACGACGACGACGACGAC 1266

QY 1266 ACCCTGGTACATCATCTGGAACCGCATGGCATGGGACATGATGATGATGATGATGAT 1325
 DB 1267 ATCCCGGCAATATAACGGGTCTATCGCATGGCATGGGATATGATGATGATGATGAT 1326

QY 1326 CGGCTACCATGATCTTGGCGTACGCGTCCGCGGTGCTCCCGGCTATATAGACATCAT 1385
 DB 1327 CGGAGCGT 1386

QY 1386 CGCGGGCTCATTTGGGGCGTCAATGTTTGGGCTTGGGCTTCTCTATGAGGAGGAGTGG 1445
 DB 1387 CTGTGTCTCAGTGGGAGTCTTGGCGGCAATAGCGTATTTCTCCATGTTGGGGAATGG 1446

QY 1446 CGAAAGT 1505
 DB 1447 CGAAGGTCTGT 1506

QY 1506 GGGGTTCTGCGCGCAGACCCCGGGCTTCCAGCGCTTATTTGACATGGGCGCCAGCG 1565
 DB 1507 GGGGAAGTGTGCGCGCGCACCGGTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1566

QY 1566 AGAAATCCAGCTCGTTAAACCAATGGCAGTGGCACAATCAACCGCACCGCCCTGAACT 1625
 DB 1567 AGAATCACTCAATGTATCAACCAACGCGAGTTGGCACAATCAATAGCAGCGCTTGA 1626

QY 1626 GCAATGACTCTTGTGACACCGGCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1685
 DB 1627 GCAATGAAAGGCTTAAACCGGCTGTTAGCAGGGCTCTTCTTATCAGCACAATTTCA 1686

QY 1686 GGTGAGGATGTCCGCAACGCAATGTCGCTGCGCGAGTATCGAGGCTTTCGGGTGGGAT 1745
 DB 1687 CTTTACGCTGTCTGAGAGGTTGGCGAGTGTGCGAGCTTACCGATTTTGGCCAGGGCT 1746

QY 1746 GGGGCGCTTGAATATGAGGATATGTCAACCAATCCAGAGGATATGAGACCTTATGCT 1805
 DB 1747 GGGGCTCTATCAGTTATGCCAA-----CGGAAGCGGCTCGACGAAACGCCCCCTACTGCT 1800

QY 1806 GGCACCTACCAAGGAGGAGTGTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1865
 DB 1801 GGCACCTACCAAGGAGGAGTGTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1860

Qy	8094	AGGTGTTCTGCGTGGAGCCCAACAGGGGGCAAGAAAGCAGCTCGCTTTATCGTTTACC	8153	Qy	9174	GGGACAGTGCAGTCCAGGCGCTCCCTCATCTCCGTGGGGGAGAGCGCGTTTGGCGTC	9233
Db	8029	AGGTGTTCTGCGTGGAGCCCAACAGGGGGCGTAAAGCAGCTCGTCTCAICGTGTTCC	8088	Db	9109	GGGCGCGGAGCGTCCGCGCTAGGCTTCTGCTCCAGAGGAGGAGGCTGCAATATGGCA	9168
Qy	8154	CTGACCTCGCGCTCAGGCTGCGGAGAAAGATGCCCTTTATGACATTTACACAAAACTTC	8213	Qy	9234	GATATCTCTTCAATTTGGCGGTGAAGACCAAGCTCAAACTCACTCACTTTCGCGAGCGC	9293
Db	8089	CCGACCTGGGCGTGGCGTGGCGAGAAAGATGCCCTGTACGAGCGTGGTTAGCAAGCTCC	8148	Db	9169	AGTACCTCTTCAACTGGGCGAGTAAAGACCAAGCTCAAACTCACTCAATAGCGCGCTG	9228
Qy	8214	CTCAGCGGTGATGGGGCTTCTTATGGATTCCAGTATCCCGCTCAGCGGTAGAGT	8273	Qy	9294	GCCTCTGATTTATTCAGCTGTTTACCGTCCGCGCGCGCGGGCGGCGACATTTATCACA	9353
Db	8149	CCCTGGCGGTGATGGGAAGCTCTTACGGAATTCGAATCTCAGGAGCAGCGGGTGAAT	8208	Db	9229	GCGGCTGAGACTTGTCCGTTGTTTACGCGCTGGCTACAGCGGGGAGACATTTATCACA	9288
Qy	8274	TTCTCTTGAAGCATGGCGGAAAGAACCTTATGGGTTTTCGTATGATACCGCAT	8333	Qy	9354	GGCTGTCGCGTCCCGACCCCGCTTATGCTCTTTGGGCGTACTCTCTTGTAGGGG	9413
Db	8209	TCTCTGTCGAAGCTGGAAGTCCAGAGAACCCGATGGGTTCTCGTATGATACCGCT	8268	Db	9289	CGGTGCTCATGCCCGCGCGCTGGTTCTGGTTTTCCTACTCTCTGCTCGCTGCAGGGG	9348
Qy	8334	GCTTTGATCAACCGTCACTGAGAGACATCAGGACTGAGGAGTCCATATATCGGGCT	8393	Qy	9414	TAGGCTTTTCTACTTCCCGCTCG--GTAGAGCGGCACACATTAGCTACACTCCATAGC	9471
Db	8269	GTTTTGACTCCACAGTCACTGAGAGCGACATCCGTACGAGGAGGCAATTTACCAATGT	8328	Db	9349	TAGGCAATCTACCTCTCCCAACCGATGAAGTTGGGGTAAACACTCCGCGCTCTTAGGC	9408
Qy	8394	GCTCTTCCCGAGGAGGCCCACTCCGATACACTCGCTAACTGAGAGACTTTAGTGG	8453	Qy	9472	TAACTGTCCTT	9531
Db	8329	GTGACCTGGACCCCAAGCGCGGTGGCCATCAAGTCCCTCACTGAGAGGCTTTATGTTG	8388	Db	9409	CATTCTCTGTTT	9468
Qy	8454	GAGGCTTATGTTCAACAGCAAGGGCCAAACCTGCGGTACAGCGTTGCGCGCCAGCG	8513	Qy	9532	TTT	9591
Db	8389	GGGCGCTCTTACCAATTCAGGGGGGAAACTGCGGCTACCGAGGTGCGCGCGAGCG	8448	Db	9469	TTCTTCT	9528
Qy	8514	GGGTGCTCAGCACTAGCATGGGAAACACATCATGCTACGTAAAGCCTTAGCGGCTT	8573	Qy	9592	ACCTTATTTTACTTTCTTCTGCTGGCTCCATCTTAGCCCTAGTCAAGCTAGCTGTGA	9651
Db	8449	GCCTACTGACAACTAGCTGTGTTAAACACCTCACTTCTGTACATCAAGGCGCGGAGCT	8508	Db	9529	TTCTTTTCTCTCTCTTTTAAATGGTGGCTCCATCTTAGCCCTAGTCAAGCTAGCTGTGA	9588
Qy	8574	GTAAGCTGAGGATTAATCGGCCCAATGCTGTTATGCGGAGTACTTGGTTGTCA	8633	Qy	9652	AAGTCCGTCGAGCGCATGACTGCAGAGAGTGCCTAACTGGTCTCTCTGCAGATCATGT	9711
Db	8509	GTGAGCGCGAGGCTCTCAGGACTGTCACCATGCTGCTGTGTGGCGACGACTTAGTGTGA	8568	Db	9589	AAGTCCGTCGAGCGCATGACTGCAGAGAGTGTGCTGATACTGGCTCTCTGCAGATCATGT	9648
Qy	8634	TCTCAGAAAGCGAGGAGCGAGGAGCGGAACCTGAGAGCTTCCACGGAGGCTA	8693	RESULT 14			
Db	8569	TCTGTGAAGTGGGGGTTCAGAGAGCGCGGAGCTTGTAGAGCTTCCACGGAGGCTA	8628	ID	AAV59364 standard; cDNA; 12980 BP.		
Qy	8694	TGACAGGTATTCCTGCTTCTGCTGAGACCCCGGAGCGGAGTATGATCTGGAGCTGA	8753	XX	AAV59364;		
Db	8629	TGACAGGTATTCCTGCTTCTGCTGAGACCCCGGAGCGGAGTATGATCTGGAGCTGA	8688	XX	11-JAN-1999 (first entry)		
Qy	8754	TAACTCTGCTCTCAATGTGCTGTGGGCTGGGCGGAGCGGCGGCGGAGTACT	8813	XX	Hepatitis C virus clone p90/HCVFLlong pu cDNA.		
Db	8689	TAACTATGCTCTCTCAACGTGTCTGCTGCGGCGGAGCGGCGGCGGAGTACT	8748	XX	HCV; therapy; diagnosis; vector; gene therapy; vaccine;		
Qy	8814	ACCTGACGAGACCTTACCTCAATCGCGGCTGCTGGGAAACAGTTAGACACT	8873	KW	p90/HCVFLlong pu; ds.		
Db	8749	ACCTTACCGTGAACCTTCAACACCCCTCGGAGAGCGGCTGGGAGACAGACACA	8808	XX	Hepatitis C virus.		
Qy	8874	CCCCGTCAATTCATGCTGGGAAACATATCCAGTACGCGCGGCGGAGCGGCTCGCA	8933	OS	WO9839031-A1.		
Db	8809	CTCCAGTCAATTCCTGCTAGGCAACATATCATGTTTGGCGGCGGAGGAGG	8868	PN	11-SEP-1998.		
Qy	8934	TGGTCTGATGACACACTTCTTCTCAATCTCATGGCTCAAGACACGCTGGACAGAAC	8993	PD	26-FEB-1998; 98WO-US04428.		
Db	8869	TGATACTGATGACCCATTTCTTATAGCGTCTCATAGGCGAGGATTCAGCTTGAACAGGCTC	8928	XX	04-MAR-1997; 97US-0811566.		
Qy	8994	TCACTTTGATGATGAGGAGCGGTACTCCGTGAGTCCCTTGGAGCTCCAGCTATAA	9053	XX	(UNIW) UNIV WASHINGTON.		
Db	8929	TTAACTGTGATCTAGGAGCGCTGCTACTCCATAGAACCTCTGATCTACTCTCAATCA	8988	XX	Kolykhalov AA, Rice CM;		
Qy	9054	TTGAAAGTTACATGGGCTTGACGCTTTTCTCTGCAACATACATATCCCGCAAGCTGA	9113	XX	WPI; 1998-520770/44.		
Db	8989	TTCAAGACTCCATGGCTCAGCGCATTTTCACTCCACAGTTACTCTCCAGGTGAATCA	9048	PT	New hepatitis C virus nucleic acid clones - comprising a 5'-terminal		
Qy	9114	CACGGTGGCTTACGCTTCAAGAAACTTGGGGGCGGCGGCGGCTCAGAGCGTGGAGGCC	9173	PT	conserved sequence, an open reading frame encoding functional		
Db	9049	ATAGGGTGGCGGCTGCTCAGAAACTTTGGGGTCCCGGCTTGGAGCTTGGAGACACC	9108	PT	components and a 3'-terminal conserved sequence		
				PS	Claim 18; Page 117-125; 209pp; English.		
				XX	This is the nucleotide sequence of a plasmid clone,		
				CC			

Db	5875	GGAAAGTCTCTCGTGGACATTTCTTCAGAGGTATGGCGGGGCTGGGGAGCTCTTGTAG	5934
Qy	5946	CATTCAAGATCATGTCTGGCGAGAGCCCTCCATGGAGATGTCGTAACCTTGTCTGCCCTG	6005
Db	5935	CCTTCAAGATCATGAGCGGTGAGTCCCTTCCACGGAGACCTGGTCAATCTGCTCCCGC	5994
Qy	6006	GAATTCTGTCTCCGGGTGCCCTTGGTAGTGGAGTCATCTGCGGGCCATCTTGGCGCGAC	6065
Db	5995	CCATCTCTCGCTGGAGCCCTTGTAGTCGGTGTCTGGCGAGCAATACTTGGCGCGC	6054
Qy	6066	ACGTGGGACCGGGGAGGCGCGCTCCAAATGAGTAATAGACTCATTTGCTTTGCTTCCA	6125
Db	6055	ACGTGCGCGCGGAGGGGAGTGAATGGATGAACCGGCTAAATAGCTTCTGCTCCCTCC	6114
Qy	6126	GAGGAATCAGCTCGCGCCCAACCACTACTGTGACGAGTCCGATGCGACGCTGTA	6185
Db	6115	GGGGGAACCATGTTTCCCCCACCACTAGCTGTGCGGAGAGCGATGACGCGCGCGTCA	6174
Qy	6186	CCCAACTACTTGGCTCCCTTACCATAACCAAGCTGTCTCAGAAGACTCCCAACTGGATTA	6245
Db	6175	CTGCCATCTCAGACGCTCACTGTAAACCACTCTCTGAGCGACTGCATCAGTGGATTA	6234
Qy	6246	CTGAGGACTGCCCATCTCCATGCGCGGCTCTGTGGCTCCGCGATGTGTGGACTGGGTTT	6305
Db	6235	GCTCGGAGTGTACCACTCCATGCTCCGGTCTCTGGCTAAGGACATCTGGGACTGGATAT	6294
Qy	6306	GCACCATCTAACAGACTTTAAAAATGGTGACCTCCAAATTAATTCCTCCAAAGATGCCG	6365
Db	6295	GCAGGTGTGAGCGACTTTAAGACCTGGCTGAAGCCAAAGCTCATGCCACAACTGCTG	6354
Qy	6366	GCCTCCCTTGTCTCTGTCAAAAGGGGTACAAAGGGCTGTGGCGCGCACTGGCATCA	6425
Db	6355	GGATTCCCTTGTCTCTGCCAGCGGGGTATAGGGGGTCTGGCGAGAGACGGCATTA	6414
Qy	6426	TGACCAACAGGTGTCTTGGCGGCGCAATATCTCTGGCAATGTCCGCTTTGGGCTCCATGA	6485
Db	6415	TGCACACTCGCTGCCACTGTGGAGCTGAGATCACTGGACATGTCAAAAACGGGACATGA	6474
Qy	6486	GAATCAGGGGCTAAGACTGTGATGAATATCTGGCAGGGGACCTTCTCTATCAATGTT	6545
Db	6475	GGATCTCGGTCTTAGGACCTCGAGAAATGTGGAGTGGGAGCTTCCCATTAACGCT	6534
Qy	6546	ACACGAGGGCCAGTGGCTGCCAAACCGCGCCAACTTTAAGTTCGCGCATCTGGAGG	6605
Db	6535	ACACGAGGGCCCTGTACTCCCTTCTTGGCGCGCACTATAAGTTCGCGCTGTGGAGG	6594
Qy	6606	TGCGGCTCTCAGAGTACGGGAGGTGACGACGACGCGGTCTACCACTACATATAACAGGAC	6665
Db	6595	TGCTCGCAGAGGAATACGTGGAGATAAGCGGGTGGGGACTTCCACTAGGTATCGGTA	6654
Qy	6666	TACCACTGATTAATCTGAAAGTCCCTGCGCACTACCTCTCTCGGAGTCTTTTCTCGGG	6725
Db	6655	TGACTACTGACATCTTAAATGCGCTGCGCAGATCCATCGCCCGAAATTTTTCACAGAT	6714
Qy	6726	TGACGAGGTGACAGATCCATAGTTTGGCCCCACACCGAGCGGTTTTCGCGGATGAGG	6785
Db	6715	TGACCGGGTGGCCCTACATAGTTTGGCCCCCTTGCAGGCGCTTGTGCGGGAGGAGG	6774
Qy	6786	TCTCTGTCTCGTGGGCTTAATTTCAATTTGTGTGGTCCGAGTCCAGCTTCTCTCGACCTG	6845
Db	6775	TATCATTCAGAGTAGACTCCAGAGTACCCGGTGGGGTCCGAATTAACCTTGGAGCCG	6834
Qy	6846	AACCCGACACAGAGTATGATGTCAATGCTAACAGATCCATCTCATATCAGCGGGAGA	6905
Db	6835	AACCGGACGTAGCGGTGTTGAGCTCCATGCTCACTGATCCCTCCCATATAACAGCAGAGG	6894
Qy	6906	CTGACGCGGGCTTTAGCGGGGGTCAACCCATCCGAGCAAGCTCTCGCGGAGCC	6965
Db	6895	CGGCGGGAGAGGTTGGCGAGAGGGTCAACCCCTTCTATGGCCAGCTCTCTCGGCAGCC	6954
Qy	6966	AGCTATCGGCACCATCGCTGCGAGCCACCTGCAACCAACCGCAAGCTTATGTTGG	7025
Db			
Db	6955	AGCTGTCCGCTCCATCTCTCAAGGCAACTTGCACGCCCAACCATGACTCCCTCGTACGCG	7014
Qy	7026	ACATGTGTGATGCTAACT-----GTTTCATGGGGCCATGTGACTCCGATAG	7073
Db	7015	AGCTCATAGAGCTAACTCTCTGTGGAGGAGAGATGGCGGCAACATCACAGGTTG	7074
Qy	7074	AGTCTGGGTCCAAAGTGGTCTCTCTCGACCCCAATGGTCCGAAGAAAGAGCG	7133
Db	7075	AGTCAGAGAACAAAGTGGTGTCTTGGACTCTCTCGATCGCTTGTGGCAGAGGAGATG	7134
Qy	7134	ACCTTGAGCTTTCGATACCATCAGAAATACATGCTCCCAAGAAAGAGTTTCCACAGCTT	7193
Db	7135	AGCGGAGGTCTCTCCGTACCGCAGAAATTTGCGGAAGTCTCGAGATTCGCGCGGCC	7194
Qy	7194	TACCGGCTGGGACAGGCTGATTAACACCCAGCCGCTTGTGGAATCGTGGAAAGGCCAG	7253
Db	7195	TGCGGCTTGGGCGCGGCGGACTACAACCCCGCTAGTAGAGACGTGGAAAAAGCCCTG	7254
Qy	7254	ATTACCAACCGGCCACTGTGTGGGCTGTGTCTCTCTCTCTAGGAAAAACCCGACGC	7313
Db	7255	ACTACGAACCACTGTGTGCTCCATGCTGCGCTACCACTCCACGGTCCCTCTCTGTC	7314
Qy	7314	CTCCCGGAGGAGCGCGACAGTGGCTTAAGTGGAGACTTCCATAGAGATGCCCTTC	7373
Db	7315	CTCCGCTCGGAAAAAGCTACCGTGTCTCTCACCGAATCAACCTATCTACTGCTTGG	7374
Qy	7374	AACAGCTGGCCATTAAGTCTTGTGGCAGCCCCCCCCCAAGCGGCGATTCAGGCTTCCA	7433
Db	7375	CGAGCTTGGCCACCAAAAGTTTGTGCACTCTCAACTTCCGGCATTAACGGGCGCAATA	7434
Qy	7434	CGGGGCGGGGCTGCGCATTCGCGAGTCAGACGCTCTCTGATGAGTTGGGCCCTTCGG	7493
Db	7435	CGACAAACATCTCTG-----AGCGCGCCCTCTCTGGTGGCCCCCGACTCG	7482
Qy	7494	AGACAGTTCATCTCTTCCATGCCCCCTCGAGGGGAGCTTGGAGATCCAGACTGG	7553
Db	7483	ACGTTGAGTCTCTTCTTCCATGCCCCCTCGAGGGGAGCTGGGATCCGATCT--	7540
Qy	7554	AGCTGAGCAGGTAGAGCCCCCAACCCCCCGGAGGGGGGTGGCAGCTCCCGCTCG	7613
Db	7541	-----CAGCGAG	7548
Qy	7614	ACTCGGGTCTCTGTCTTCTCGAGGAGGAGCTCCGCTGCTGTGCTCTCATGT	7673
Db	7549	GGTCATGTGCAAGCTGAGTAGTGGGCGGACACGGAAGATGTCTGTCTCTCAATGT	7608
Qy	7674	CATCTCTGACGCGGGCTCTAATACTCTTGTAGTCCGAGAGAGAGTTACCGA	7733
Db	7609	CTTATCTCGACAGCGCACTCTGTCACCCCGTGGCTGCGGAGAACAAAACTGCCA	7668
Qy	7734	TTAACCCCTTGACNACTCTCTGTCGATATCACAAAGTGTACTGTACCAACNA	7793
Db	7669	TCAACGCACTGAGCAACTCTGTGTACGCCATCAATCTGGTGTATTTCCACCACTCAC	7728
Qy	7794	AGAGCGCTCACTAAGGCTTAAAGGTAACTTTTGTAGATGCAAGTCTCGACTCT	7853
Db	7729	GCAGTCTTGCAGAGGAGAGAAAGTCAATTTGACAGACTGCAAGTCTTGGAGGCC	7788
Qy	7854	ACTACGACTCAGTCTTAAAGGACATTAAGCTAGCGGCTCCAAAGTCCCGCAAGGCTCC	7913
Db	7789	ATTACGAGCGTCTCAAGGAGGTCAAGCAGCGGCTCAAAAGTGAAGGCTAATCTGC	7848
Qy	7914	TCACCATGGAGGCTTGGCAGTTAAACCCACCCCATTTCTGCAAGATCTAATATGGT	7973
Db	7849	TATCCGTAGAGGAAGCTTGCAGCCTGACGCCCCCACTTCAGCCAAATTCGAAGTTGCT	7908
Qy	7974	TTGGGCTTAAAGAGTCTCGCAGCTTGTCCGGAGGGCGGCTTAAACCAATCAAGTCCGTGT	8033
Db	7909	ATGGGCAAAAGACGTCTGTTGCCATGCCAAGAGGCGGTAGCCCACTCACTCTCGTGT	7968
Qy	8034	GGAGGACCTCTCGAGGACTCAGAAACAAATTTCCCAACCAATTTATGCGCAAAATG	8093
Db	7969	GGAAAGACCTTCTGGAAGACAGTGTAAACCAATAGACACTACCATCATGCGCAAGACG	8028

Qy	3726	GTGGAGCGGTCGACCTGTACCTGCTCAGCGGAAACGCTGATGTCATCCCGGCTCGAAGAC	3785	4806	GCACGCGCGGGTCCGACCGGTAGGGGAAGACTGGGCATTTATAGTATGTTTTCACCTG	4865
Db	3715	CGCGCTCTCGGACCTTACCTGGTCAGAGGACGCGGATGTCATTCGCTGGCGCGC	3774	4795	CTCAACGCGCGGAGGACTGGCAGGGGAAGCCAGGCATCTACAGATTTGTGGCACCGC	4854
Qy	3786	CGCGGAGCAAAACGGGAGCGCTACTCTCCCGAGAGCTCTTTTCACTTGAAGGGTCTCT	3845	4866	GTGAGGAGCCTCAGGAATTTTTCAGAGTGTAGTCTCTGCTGCTAGTGTCTAGCAGCAGGGG	4925
Db	3775	GAGGTATAGCAGGGGTAGCTGCTTTTCGCGCGGCGCAATTTCTACTTTGAAGGCTCTCT	3834	4855	GGGAGCGCCCTTCGCGCATGTTTCGACTCTGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCT	4914
Qy	3846	CAGAGGCGCGGTCTATGCCCCCAGGGCCAGCGCTGTCGAGTCTTCCGCGCAGCTGTGT	3905	4926	CGCGATGCTATGAGCTCACAACCATCGGAGACCAACCGCTCAGGCTCAGGGCTATTTTCAACA	4985
Db	3835	CGGGGGTCCGCTGTGTGCCCCCGCGGACACGCGTGGGCTTATTCAGGGCGCGGCTGT	3894	4915	GTGCTTGTGTATGAGCTCAGCGCGCGAGACTACAGTGTAGGCTACGAGCTACATGAACA	4974
Qy	3906	GCTCTCGGGCGGTGAGTCAATAGATTTTATCCCGCTTGAGACACTCGACATCGTCA	3965	4986	CGCGCGGTTTGGCTGTGTGCCAAGACCATCTTGTAGTTTTTGGAGGAGGTTTTTTCACCGGC	5045
Db	3895	GCACCGTGTGAGTGGCTTAAGCGGTGACTTTATCCCTGTGGAGACCTTAGACACAACA	3954	4975	CCCGCGGGCTCCCGGTGCGCAGGACCATCTTGAATTTTGGAGGGCGCTTTTACGGGC	5034
Qy	3966	CGCGTCCCCACCTTTTGTGACAAACAGACACACCTGTGTGTGCCCGAGACTATCAGG	4025	5046	TCACACACATAGATGCGCACTTCTTCCCAACAAGCAATCGGGGGAATTTTCGCAT	5105
Db	3955	TGAGATCCCGGCTGTTTCAAGGACAACTCTCTCCACGAGCTGCCCGAGCTTCCAGG	4014	5035	TCACCTCATATAGATGCGCACTTCTTATCCCAACAAGCAGAGTGGGGAGAACTTTCTCT	5094
Qy	4026	TCGGGTACTTGCATGCCCGACTGGCAGTGGAGAGGACCAAGTTCTCTGTCGCATATG	4085	5106	ACTTAAACGCTTACAGGCTTACAGTGTGCGCTAGGGCCAAAGCCCCCCCCCTCTCTGGG	5165
Db	4015	TGGCCCCACCTGCTGCTCCACCGGAGCGGTAGAGACCAAGGTCGCGCTGCGTAGC	4074	5095	ACTGTGTAGGCTTACCAAGCCACCGTGTGCTAGGCTCAAGCCCTCCCGCATCTGGG	5154
Qy	4086	CTGCTCAGGGGTATAAAGTGTAGTCTTAATCCCTCAGTGGCTGCCACCCCTCGGGTTTG	4145	5166	AGCTCATGTGGAGTGTGTGACTCGACTCAAGCCCACTGTGTGGGGCCCACTCTCTCC	5225
Db	4075	CAGCCAGGCTTACAGGTGTGTGTCTCAACCCCTCTGTGTGCAACCTCGGGCTTTG	4134	5155	ACCAGATGTGGAAGTGTGTGATCGGCTTAAACCCCACTCTCCATGGGCGCAACCCCTGC	5214
Qy	4146	GGGGTACTTGTCTAAGGCAATGGCAATCAATCCCAATTTAGGACTGTGAGTCAAGACTG	4205	5226	TGTACCGCTTGGCTGTGTTCACAGAGGTTCACCTCACAATCCCGTGAGCAATATACA	5285
Db	4135	GTGCTTACATGTCCAGGCTCATGGGTTGATCTTAATATCAGAGCCGGGTGAGAACAA	4194	5215	TATACAGACTGGGCGCTGTTCAGATGAAGTCAACCTGACGCAACCCCAATCACCAATACA	5274
Qy	4206	TGACGCGGGGCGCCATCAGCTACTCCACATATGCAATATGCAATTCCTCGCGATGGGGCT	4265	5286	TCGCCACCTGCAATGCAAGCCGACCTTGTAGGTATGATCAGCAGCATATGGGTCTTGGCAGGG	5345
Db	4195	TTACCACTGGCAGCGCCATCAGCTACTCCACCTACGGCAAGTTCCTTGGCGGAGCGGGT	4254	5275	TCATGACATGCACTGTGCGCGGACCTGAGGTGCTCAGAGCACTTGGTCTCTGTGGCG	5334
Qy	4266	GTGCGGCGCGGCTACGACATCATATATGTATGATGATGCCATGCGCTGGACTCTACCA	4325	5346	GAGTCTTGGCGCGCTGCGCGGTATGCTGTGGCGGACCGGGTGTGTGTGATCATCGGCC	5405
Db	4255	GCTCAGAGGTGCTTATGACATAATATTTGTGACGAGTGCCTCCACGAGTCCACAT	4314	5335	GCCTCTGCTGCTGTCTGCGCGGCTATTTGCTGTCTCAACAGGCTGCTGTATAGTGGCA	5394
Qy	4326	CCATCTTGGGATCGGAACAGTCTTGATCAAGCAGAGACAGCTGTGGGTGAGACTAATCTG	4385	5406	GCTTGCACATTAACACAGCAGCGCTGCTGCGCGGCAAGAGAGTCTCTATAGAGCTT	5465
Db	4315	CCATCTTGGGATCGGCACTGTCTTACCAAGCAGAGACTGCGGGGCGGAGACTGTGTG	4374	5395	GGATTTGCTTGTCCGGGAAAGCGGCAATATATCTGACAGGGAGGTCTCTACAGAGT	5454
Qy	4386	TGCTGGCTACAGCTACGCGCCCTGGGTTCAGTGACAAACCCCGCCCAACATAGAGGAGG	4445	5466	TTGATGAGATGGAGGAATGTGCTCTAGGGCGGCTCTCTAGGGCGGCTCTCTAGAGGAGG	5525
Db	4375	TGCTCGCCACTGTCTACCCCTCCGGGCTCCGCTACTGTCTCCCATCTTAAACATCGAGGAG	4434	5455	TCGATGAGATGGAGAGTGTCTCTCAGCACTTACCGTACATCGAGCAAGGATGATGCTCG	5514
Qy	4446	TGGCCCTTGGCAGGAGGCGAGATCCCTTCTATGGAGGGGATTCCTCTGTCTTACA	4505	5526	CCGAGATGTGAAAGTCCAAAGATCCAAAGGCTTATTGAGCAAGCTTCCAAACAGCTCAAG	5585
Db	4435	TTGCTCTGTCCACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCTCGAGGTGA	4494	5515	CTGAGCAGTTCAAGCAGAGAGGCGCTCGGCTCTCTGAGACCGGCTCCCGCAAGCAGAGG	5574
Qy	4506	TCAGGAGGAGAGACATCTGATCTTCTGCCATTCAAAGAAAGTGTGACGAGCTCGCGG	4565	5586	AGATACAAACCATGTGCGAGGCTTCAATGCGCCCAAGGTAGAACAAATTTCTGGGCCAAGACA	5645
Db	4495	TCAGGAGGAGAGACATCTCATCTCTGCGCACTCAAGAGAGAGTGTGCGAGCTCGCGG	4554	5575	TTATCAACCCCTGTCTGTCAGACCAACTGGGAGAACTCGAGGTCTTCTGGCGGAGAGACA	5634
Qy	4566	CGGCGCTTGGGGTATGGCTTGAATCTCAGTGGCACTACAGAGGTTTGGAGTCTCGG	4625	5646	TGTGGAATCTCATTTAGCGGATCCAAATACCTCCGAGGACTTATCAACACTTCCAGGGAACC	5705
Db	4555	CGAAGCTGTGGCATTTGGGCATCAATGCGCGTGGCTTACTACCGGGTCTTGAAGTGTCTG	4614	5635	TGTGGAATTTCACTCAGTGGGATACATACTTGGCGGCGCTGTCAACGCTGCTGTTAACC	5694
Qy	4626	TAATACCAACTCAGGAGAGAGTGTGTGTGCTGCGCAGCGGCTCATGACAGGGTATA	4685	5706	CTCAGTGTGCTTCCATGATGGCGCTTCAAGTCCCGCTCTACAGTCCGCTGTCAACAGCA	5765
Db	4615	TCATCCCGACCGCGGAGTGTGTGTGCTGCTGCGAGATGCTCTCATGACTGGCTTTA	4674	5695	CCGCCATTTGCTTCAATGATGGCTTTTACAGTCTGCGCTCACAGCCCACTAACCACTGGCC	5754
Qy	4686	CTGGGAGCTTGTACTCCGCTGATCGACTGCAACAGTGTGGGTGCTCAAGTGTGTAGACTTCA	4745	5766	CCACTATCTTCTCAACATTTTGGGGGCTGGCTAGCATCCCAAAATTTGACCAACCCCGCGG	5825
Db	4675	CCGCGGACTTGCATCTGTGTAGACTGCAACAGTGTGTCACTCAGACAGTGTGATTTCA	4734	5755	AAACCCCTCTCTTCAACATATTTGGGGGGTGGGTGGCTGCGGAGCTCGCGCCCGCGGTG	5814
Qy	4746	GTTTAGACCCCAATTACCATATACCAACAGATTTGCTTCCCTCAAGACGCTGTCTCAGTA	4805	5826	GGGCGACTGGCTTGTGTGTGCTCAGTGGGCTAGTGTGGAGTGTGCTGAGGAGTATAGGCTTAG	5885
Db	4735	GCCTTGACCTTACCTTTTACCATATGAGACAACACGCTCCCGCCAGGATGCTGTCTCCAGGA	4794	5815	CCGCTACCGGCTTGTGGGCGCTGGCTTGTAGTGTGCGCGCGGCTATCGGAGCGGTGGACTGG	5874
Qy				5886	GTAAGGTGTGTAGTGGACATCTCTGGCAGGGGTATGGTGGGGCAATTTTGGGGGGCTCTCTCGTCG	5945

1507 GGGGAAGTGGCGCGCACACAGCTGGGCTTGTGTCTCTCTTACACAGGCGCCAAAGC 1566
1566 AGAAATCCAGCTCGTTAAACCAATGGAGCTGGACATCAACCGCACCGCCCTGAACT 1625
1567 AGAAATCCAACTGATCAACCAACAGCGAGTGGACATCAATAGACAGCGCTTGAAC 1626
1626 GCAATGACTCTTGCACACCGGCTTATCGGCTCTGTGTTCTACACCAAGCTTCAACT 1685
1627 GCAATGAAGCTTAAACCGGCTGGTTAGAGGGCTCTTCTATCAGCACAATTTCACT 1686
1686 CGTCAGAGTTCGCAACGATGTCGCCCTGCCGAGATCGAGGCTTCCGGGTGGGAT 1745
1687 CTTGAGGCTCTCTGAGAGGTGGCCAGCTGCCGAGCTTACCGATTTTGGCCAGGCT 1746
1746 GGGGCGCTTGCATATAGAGTAATGTACCAATCCAGAGGATATGAGACCTTATGCT 1805
1747 GGGGTCTTATCAGTTATGCCAA-----CGGAAGCGGCTCTGACGAAACGCCCCCTACTGCT 1800
1806 GGCACTACCCACCAAGGAGCTGGGCTGTCTCCGGAAGACTGTGTGTGGGCCCAAGTGT 1865
1801 GGCACTACCTCCAGACCTTGTGGCATTTGCCCGCAAGAGCGTGTGTGGCCCGGTAT 1860
1866 ACTGTTTACCCCGACAGTGTGTGGGACGACCGACAGGCTTGGAGCGCCCACTT 1925
1861 ATTGCTTCACTCCAGCGCGTGTGTGGGAACGACCGACAGGCTCGGGCGCGCTTACCT 1920
1926 ACAGTGGGGGAGATGACAGAGTGTCTTCTATTTGAACAGCACTCGACACCGCTGG 1985
1921 ACAGTGGGGTGAATGATACGGATGTCTTCTTAAACACAGGCGCACCGCTGG 1980
1986 GGTTCATGTTTCGGCTGACGCTGATGAATCTTCTGTGCTACACCAAGACTTGGCGGCGAC 2045
1981 GCAATGGTTCGGTTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
2046 CACCTTCCGCTTACGAGTGTCTTAAACCGGACGAGGCTGTGTGTGGCCCGACGAGCT 2105
2041 CCCCTTGTGTCTCGGAGGCTGGGCAACAAC-----CCTTGTCTGCGCCCACTGAT 2094
2106 GTTTAGGAGATCTGATACCACTTACCTCAATGCGGCTCTGGGCTCTGCTCAGCG 2165
2095 GTTTCCGACAGCATCCGGAAGCACATCTCTCGGTGCGGCTCCGCTCGGATTAACAC 2154
2166 CAAGTGCCTGATGCTACCTTACAGGCTCTGGCAATTAACCTGACACAGTGTAACTATA 2225
2155 CGAGTGATGTTGCTACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2214
2226 CCATCTTCAAAATAAGGATGATGTGGAGGGGTGAGCAGGCTCAGCGCTGCAATGCA 2285
2215 CCATATTTCAAGTCAGGATGTAGTGGAGGGGTGAGCAGGCTGGAAGCGGCTGCA 2274
2286 ATTTCACTCTGGGATCGTTGCACTTGGAGACAGAGCAGAGTCAACTGTCTCTT 2345
2275 ACTGGACGCGGGGCAACGCTGTGATCTGGAAGACAGGACAGGCTCGAGCTCAGGCCAT 2334
2346 TGTGCACTCCACAGGAATGGCCATTTTACTTCTCTTACTCGGACTCGGCTCGGCT 2405
2335 TGCTGCTGTCACACAGTGGAGGCTCTTCTGCTGTTCTTTTCAAGCCCTGCAAGCT 2394
2406 TGTGCACTGCTTCTCACCTCCACAAACATCGTGGAGCTGATCAATTCATGTATGGCC 2465
2395 TGTCCACCGGCTCATCCACTCCACAGAACATTTGGAGCGTGCAGTACTTGTACGGGG 2454
2466 TATCACTGCTTCTCACAAATATCATCGTCCGATGGAGTGGGTAATCTTATCTCTGTC 2525
2455 TAGGGTCAAGCATCGGCTCTCGGCTTAAAGTGGGAGTACGCTGTTCTCTCTCTCTCC 2514
2526 TCTTAGCGAGCGCAGGCTTTCGCTGCTTATGATGCTCATCTTGTGGGCGACGCG 2585
2515 TGCTTGAGACGCGGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2574
2586 AAGCAGCACTAGAGAAGCTGGTCACTTTGACGCTGCGAGCGAGCTAGCTGCAATGGCT 2645

2575 AGCGCGCTTTGGAGAACCTCGTAATACTCAATGCAAGCTCCCTGGCGGAGCAGCGTC 2634
2646 TCTTATTTTGTGATCTTTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2705
2635 TTGCTGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2694
2706 TAGCTACCTTATTTCCCTCACTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2765
2695 GAGCGGTCTACGCTTCTACGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2754
2766 AACAGCTTATGCTTATGACGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2825
2755 AGCGGCGATACGCACTGACAGCGAGGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2814
2826 TGATCACTCTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2885
2815 GGTTAATGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2874
2886 GGTGCTGCTTCTTCTGACCTTGGGGAGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2945
2875 GGTTCAGTATTTCTGACCGAGTAGAAGCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2934
2946 AGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3005
2935 AGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2994
3006 TGTGACATAACCAAGTGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3065
2995 TATTTGACATACCAAACTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3054
3066 CTTTGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3125
3055 GTTGTCTTAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3114
3126 CAAGGATCTCGCGGGGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3185
3115 CGCGAAGATAGCGGAGTCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3174
3186 CTGCACTTACATCTATGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3245
3175 CTGCACTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3234
3246 GGGACTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3305
3235 GAGATCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3294
3306 TCTGGGAGCGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3365
3295 CGTGGGGGAGATACCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3354
3366 GACTTGGTGGGAGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3425
3355 GTAGGGGCGAGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3414
3426 TTTCTGCGCCCACTGCTTACGCGCCAGCAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3485
3415 TGCTGCGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3474
3486 TGAGCATGAGCGGGGCGGAG 3545
3475 CCAGCTTGAATGCTGCGGGGAG 3534
3546 TCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3605
3535 CTACCCAAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3594
3606 CTGCGAAACAAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3665
3595 CGGGAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3654
3666 AGGGGAGCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3725
3655 ACCAAGACCTTGTGGGCTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3714

KW Hepatitis C virus; HCV; productive replication; infection; antiviral;
 KW plasmid p90/HCVFlongpu; ss.
 XX
 OS Hepatitis C virus.
 XX Synthetic.
 XX
 PN US6392028-B1.
 XX
 PD 21-MAY-2002.
 XX
 PF 04-MAR-1998; 98US-0034756.
 XX
 PR 04-MAR-1997; 97US-039843P.
 PR 04-MAR-1997; 97US-0811566.
 XX
 XX (UNIW) UNIV WASHINGTON.
 PA
 XX Rice CM, Kolykhalov AA;
 PI
 XX WPI; 2002-478540/51.
 DR
 XX
 XX New DNA or RNA comprising a hepatitis C virus sequence that contains 3'
 PT and 5' non-translated regions flanking a polyprotein coding region,
 PT useful for detecting and developing treatment for infection by
 PT hepatitis C virus -
 PT
 XX
 XX Example 3; Column 89-102; 114pp; English.
 PS
 XX The present invention relates to a new DNA or RNA comprising a
 CC hepatitis C virus (HCV) sequence capable of productive replication
 CC in a host cell. The molecules of the invention comprise 5' to 3' on
 CC the positive sense strand a functional HCV 5' non-translated region
 CC comprising an extreme 5'-terminal conserved sequence, an HCV polyprotein
 CC coding region and a functional HCV 3' non-translated region comprising
 CC an extreme 3'-terminal conserved region. The DNA or RNA is used to
 CC detect and develop treatment for infection by hepatitis C virus. The
 CC present nucleic acid sequence represents the hepatitis C virus plasmid
 CC p90/HCVFlongpu cDNA sequence.
 XX
 XX Sequence 12980 BP; 2593 A; 3860 C; 3632 G; 2895 T; 0 other;
 SQ
 Query Match 47.6%; Score 4621.6; DB 24; Length 12980;
 Best Local Similarity 67.8%; Pred. No. 0;
 Matches 6594; Conservative 0; Mismatches 3034; Indels 92; Gaps 6;
 6 CCCTAATAGGGCGGACACTCCGCATGAATCACTCCCTCTGAGGAACACTGTCTTCA 65
 7 CCCTGTAGTGGGGCGACACTCCACATGAATCACTCCCTCTGAGGAACACTGTCTTCA 66
 66 CGCAGAAAGCGTCTAGCCATGGCGTGTAGTATGAGTGTCTGATAGCTTCCAGGCCCCCC 125
 67 CGCAGAAAGCGTCTAGCCATGGCGTGTAGTATGAGTGTCTGATAGCTTCCAGGCCCCCC 126
 126 TCCCGGAGAGCCATAGTGTCTCGGAAACCGGTGAGTACACCGGAATGCCCGGAAAGAC 185
 127 TCCCGGAGAGCCATAGTGTCTCGGAAACCGGTGAGTACACCGGAATGCCAGGACGAC 186
 186 TGGTCTCTTCTTGATTAACCCACTCTATGCGCGGCCATTTGGCGGTGCCCGGAAAGA 245
 187 CGGGTCTCTTCTTGATTAACCCACTCTATGCGCGGCCATTTGGCGGTGCCCGGAAAGA 246
 246 CTGCTAGCCGAGTAGCGTTGGGTTGCGAAAGGCCCTGTGGTACTGTGCTGATAGGGTCTT 305
 247 CTGCTAGCCGAGTAGCGTTGGGTTGCGAAAGGCCCTGTGGTACTGTGCTGATAGGGTCTT 306
 306 GCGAGTGGCCCGGAGGTCTCGTAGACCGTGTGATAGTACACCAAAATCTTAAACCTCAAA 365
 307 GCGAGTGGCCCGGAGGTCTCGTAGACCGTGTGATAGTACACCAAAATCTTAAACCTCAAA 366
 366 GAAAAACCAAAAGAAACACCAACCGTGCCTGCCCAAGAGGTTAAGTTTCCGGGCGGCGCC 425
 367 GAAAAACCAAAAGAAACACCAACCGTGCCTGCCCAAGAGGTTAAGTTTCCGGGCGGCGTC 426

QY 426 AGATCGTTGGCGGAGTATATCTTTGTTCCCGCGAGGGGCCCCAGGTTGGTGTGCGCGCA 485
 DB 427 AGATCGTTGGTGGAGTTTACTTTGTTGGCGCGAGGGGCCCTAGATTGGGTGTGCGCGCA 486
 QY 486 CAAGGAGACTTTCGAGCGGTCCAGCCAGCTGGAAGGCGCGAGCCCATCCCTAAAGATC 545
 DB 487 CGAGGAAGACTTTCGAGCGGTCCCAACCTCGAGGTAGACGTGAGCCATATCCCAAGGCAC 546
 QY 546 GCGCTCCACTGCGCAAACTCTGGGGAACACAGGATACCCCTGGGCCCTTATACGGGAATG 605
 DB 547 GTGCGCCGAGGCGAGGACTTGGGCTACGCCCGGTACCTTGGGCCCTCTATGGCAATG 606
 QY 606 AGGAGCTCGGCTGGGAGGATGGCTCTGTCTCCCGCGAGGTTCCTGCTCTTTGGGGCC 665
 DB 607 AGGGTTGCGGGTGGCGGATGGCTCTGTCTCCCGTGGCTCTCGGCCCTAGCTGGGGCC 666
 QY 666 CCNATGACCCCGGCGATAGTGTGCGCAACGTGGGTAAAGTTCATGATACCTTAACGTGG 725
 DB 667 CCACAGACCCCGGCGTAGTTCGCGCAATTTGGGTAAAGTTCATGATACCTTACGTGG 726
 QY 726 GCTTTGCCGACCTCATGGGGTATACCTCTGTGTGGGCGGCCCGCTCGGCGGCTCGCCA 785
 DB 727 GCTTCCGCGACTCATGGGGTATACACCGCTGTGCGCGGCCCTCTTGGAGGCGCTGCCA 786
 QY 786 GAGCTCTCGCGCATGCGGTGAGAGTCTCGAGAGACGGGGTTAATTTTGCACAGGGAAT 845
 DB 787 GGGCCCTGGCGCATGGCGTCCGGTTCTGGAAGACGGCGTGAATATGCAACAGGGAACC 846
 QY 846 TACCGGTTGCTCTTTTCTATCTTCTGCTGGCCCTGCTGCTGCTCATCACACCCCG 905
 DB 847 TTCTGCTGGTCTCTTTCTATCTTCTGCTGGCCCTGCTCTCTGCTGACCGTGGCCG 906
 QY 906 TCTCCGCTCCGAAGTGAAGAACATCAGTACCGGCTACATGGTGAATTAACGACTGCACCA 965
 DB 907 CTTGACCTACCAAGTGGCAATTCCTCGGGGCTTTACCATGTACCAATGATGGCTTA 966
 QY 966 ATGACAGCATTAACCTGGAGTCCAGGCTGCTGTCTCTCAACGTCGCGGGTGGTCCCGT 1025
 DB 967 ACTCGAGTATTTGTGTACGAGGCGCGGATGCTCATCTCCGACACTCCGGGGTGTCTC 1026
 QY 1026 CGGAGAAAGTGGGAATGATCTCAGTGTGGATACGGGCTCACCAGATGTGCGCGCTGC 1085
 DB 1027 GCGTTCGCGAGGGTAAAGGCTCGAGGTGTGGGTGGCGGTGACCCCGACGGTGGCCACCA 1086
 QY 1086 AGCGGCGCGCGCCCTCAGCGAGGGCTTGGCGAGCAGCATCGACATGTTGTGATGTCCG 1145
 DB 1087 GGGACGGCAAACTCCCCACAAACGAGCTTCGAGCTCATATCGATCTGCTTGTGCGGAG 1146
 QY 1146 CCACGCTCTGCTGCGCCCTCTACGTGGGGACCTGTGCGGGTCTGCTTTCTTTGTGGTC 1206
 DB 1206 AATGTTTCACTGCTCGCGCAGCAGCACTGGTTTGTCCAAAGACTGCAATGCTCCATCT 1265
 DB 1207 AACTGTTTCACTTCTTCCAGCGCCACTTGAACGAGCAGCAATGTTGTCTATCT 1266
 QY 1266 ACCCTGTACATCACTGGACACCGCATGGGATGGGACATGATGATGATGCTGCGCCA 1325
 DB 1267 ATCCCGCCCATTAACGGGTTCATCGCATGGCATGGATGATGATGATGATGATGATG 1326
 QY 1326 CGGCTACCATGATCTTTGGCGTACCGATCGGTGCTCCCGAGGTCAATTTATAGACATATTA 1385
 DB 1327 CGGCGAGGTTGGTGGTAGCTCAGCTGCTCCGATCCCAAGCATCATGGACATGATCG 1386
 QY 1386 GCGGGGCTCATTTGGGGGTCATGTTGGCTGGCTCTTCTCTATGCGAGGAGCGTGGG 1445
 DB 1387 CTGGTGTCACTGGGGAGTCTGCGCGGCATAGCTGATTTCTCATGGTGGGGAACCTGGG 1446
 QY 1446 CGAAGTCTGTTGTCATCTTCTGCGCGCGGGTGGAGCGCGCAGCCATCTACTCTGTTG 1505
 DB 1447 CGAAGTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1506
 QY 1506 GGGGTTCTGCGCGCAGACCAACCGGGCGCTTCAACAGCTTTATTTGACATGGGCCCCAGGC 1565

Db 7609 CTTATTCTGGACAGGCGCACTCGTCACCCCGTGGCTGGCGAAGAACAAAACACTGCCCA 7668
 Qy 7734 TTAACCCCTTGGACCACTCCTGTTGCGATATCACACAAAGTGTACTGTACCAACAA 7793
 Db 7669 TCAACGCACTGAGCACTCGTTGCTAGCGCATCACAACTGGTGTATTCACCACTTCAC 7728
 Qy 7794 AGACGCGCTCACTAAAGGGCTAAAAGGTAACTTTTGTATAGATGCAAGTGTCTCGACTCCT 7853
 Db 7729 GCAGTGTCTGCCAAAGCAGAGAAAGTCACTTTGACAGACTGCAAGTTCTGGACAGCC 7788
 Qy 7854 ACTACGACTCAGCTTTAAAGGACATTAAGCTTAGCGGCTTCAAGGTCACCGCAAGGCTCC 7913
 Db 7789 ATTACCAAGGACGTCTCAAGGAGCTCAAGAGAGCTCAAGAGCGCGCTCAAAAGTCAAGGCTTAACTTGC 7848
 Qy 7914 TCACCATGAGGAGGCTTCCAGTTAACCCACCCCACTTCTGCAAGATCTAAATATGGGT 7973
 Db 7849 TATCCGTAGAGGAGCTTGCAGCTGACGCCCCCACTTCCAGCCAAATCCAAAGTTGGCT 7908
 Qy 7974 TTGGGGCTAAGGAGGCTCCGAGCTTGTCCGGGAGGCGGTTAAACCATCAAGTCCGCTGT 8033
 Db 7909 ATGGGGCAAGAGCTCGTTGCTGATGCCAAGAGGCGGTAGGCCCACTCACTCCGCTGT 7968
 Qy 8034 GGAAGGACCTCTGGAGGACTCAGAAACACCAATTTCCCAACCACTTATGGCCCAAAATG 8093
 Db 7969 GGAAGGACCTCTGGAAGACAGTGAACCAATAGACACTACCATCATGCGCCCAAGACG 8028
 Qy 8094 AGGTGTTCTCGTGGAGCCCAACAGAGGGGGGCAAGAAAGAGCTCGCTTATCGTTTACC 8153
 Db 8029 AGGTTTTCTCGTTACGCTGAGAAAGGGGGGCTGTAAGCCAGCTCGTCTCATCGTGTTC 8088
 Qy 8154 CTGACCTCGCGCTCAGGCTCGAGAGAGATGCCCTTTATGACATTTACACAAAATTC 8213
 Db 8089 CCGACCTGGCGCTGCGGTGTGAGAGAGATGSCCTTTAGAGCTGTTAGCAAGCTCC 8148
 Qy 8214 CTCAGGCGGTGATGGGGGCTTCTTATGGATTCAGATATCCCGCGCTCAGCGGGTAGAGT 8273
 Db 8149 CCCTGGCGGTGATGGGAAGCTCTACGATTCCTCAATACCTCACCAGGACAGCGGGTTGAAT 8208
 Qy 8274 TTCTCTGAAAGCATGGCGGGAAGAAAGAGCCCTATGGTGTGTTGATGATACCGGAT 8333
 Db 8209 TCCTCTGTCAGCGTGAAGTTCAGAAAGACCCCGATGGGGTTCCTGATGATACCCGCT 8268
 Qy 8334 GTTTGACTCAACCGTCACTGAGAGAGACATCAGGACTGAGGAGTCCATATATCGGGCCT 8393
 Db 8269 GTTTGACTCCAGTCACTGAGAGCGACATCCGTACGGAGGAGGCAATTTACCAATGTT 8328
 Qy 8394 GCTCCTTGGCCGAGGAGGCCACACTGCCATACACTCGCTAACCTGAGAGACTTTACGTGG 8453
 Db 8329 GTGACCTGGACCCCAAGCCCGGTGGCCATCAAGTCCCTCACTGAGAGGCTTTATGTTG 8388
 Qy 8454 GAGGCGCTATGTTCAACAGCAAGGCGCAACCTGCGGGTACAGCGGTGCGCGCCAGCG 8513
 Db 8389 GGGGCGCTCTTACCAATTCAGGGGGGGAAGTTCGCGCTACCCGAGGTGCGCGAGCG 8448
 Qy 8514 GGGTGTCCACCTAGCATGGGAACACCATCATGCTACGTAAGGAAAGCTTACGCGCTT 8573
 Db 8449 GCGTACTGCAACTAGCTGTGGTAAACACCTCACTTGTCTACATCAAGGCGCCGGAGCGCT 8508
 Qy 8574 GTAAAGCTGAGGAGTAATCCGCGCCCAATGTGTTGATGCGGCGATGACTTGGTTGTCA 8633
 Db 8509 GTCAGGCGGAGGCTCCAGACTGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8568
 Qy 8634 TCTCAGAAAGCAGGAGCCGAGAGGAGCAGAGGAGCAGAGGAGCCTGAGAGGCTTACCGAGGCTA 8693
 Db 8569 TCTGTGAAAGTGGCGGGGTCCAGAGGAGCAGCGCGAGCTTACGAGGCTTACCGAGGCTA 8628
 Qy 8694 TGACCAAGTATCTGCGCCCTCTGCTGAGCCCGCCAGAGCCGAGATGATCTCGAGCTGA 8753
 Db 8629 TGACCAAGTATCTGCGCCCGCCCGGGGACCCCGCCAGAGGAGTACGATCTGAGGCTTA 8688
 Qy 8754 TAACATCTTCTCTCAAAATGTGTCTGTGGCGTGGGCGCCCAAGGCGCGCCAGATACT 8813

Db 8689 TAACATCATGTCTCTCCAAACGTGTGAGTGGCCACGACGGCGCTGGAAGAGGGTCTACT 8748
 Qy 8814 ACCTGACAGAGACACCTACCTACCAATCGCCCGGCTGCTGGGAACAGTTAGACACT 8873
 Db 8749 ACCTTACCGGTGACCTTCAACCCCTCGCAGAGAGCCGCTGGAGACAGCAACACCA 8808
 Qy 8874 CCCCTGTCAATTCATGGCTGGGAACATCATCAGTACGCCCGGACCAATATGGGCTCGCA 8933
 Db 8809 CTCAGTCAATTCCTGGCTAGGCAACATAATCATGTTTGGCCCCACACTGTGGGCGAGCA 8868
 Qy 8934 TGGTCTGTATGACACACTTCTCTCATGCTCATGCTCATGCTCATGCTCATGCTCATGCT 8993
 Db 8869 TGATATCTGATGACCACTTCTTATAGCTCTCATGACCCAGGATCAGCTTGAACAGGCTC 8928
 Qy 8994 TCAACTTTGAGATGTACGAGCGGTGTACTCCGTGAGTCCCTTGGACCTCCCGAGCTATAA 9053
 Db 8929 TTAACCTGTGAGATCTACGAGGCTGCTACTCTCATGAGACCTTGAATCTACTTCCATCA 8988
 Qy 9054 TTGAAAGGTTTACATGGGCTTGAACGCTTTTCTCTGTCACACATACACTTCCCGACGAACTGA 9113
 Db 8989 TTTCAAGACTCCATGSCCTCAGCGCATTTTCACTCCACAGTTACTCTCCAGGTGAATCA 9048
 Qy 9114 CACGGTGGCTTTCAGCCCTCAGAAACCTTTGGGGGCGCACCCCTCAGAGGCTGGAAGAGCC 9173
 Db 9049 ATAGGGTGGCGCATGCTCTCAGAAACCTTTGGGGTCCCGCTTGGAGCTTGGAGACACC 9108
 Qy 9174 GGGCAGTGCAGTTCAGGGGCTCCCTCATCTCCGTTGGGGGAGAGCGGCTTGGGCTC 9233
 Db 9109 GGGCGCGGAGCGTCCGCGCTAGGCTTCTGTCCAGAGGAGGAGGCTGCGCATATGTGGCA 9168
 Qy 9234 GATATCTCTTCAATTTGGGCGGTGAAGACCAAGCTCAAACTCACTCCATTTGCCGGAAGCCG 9293
 Db 9169 AGTACTCTTCACTGGGCGAGTAAAGCAAGCTCAAACTCACTCCATAGCGCGCGCTG 9228
 Qy 9294 GCTCCTGTGATTTATTCAGCTGGTTTACCGTCCGCGCGCGCGGCGGCGAGCAATTTATCA 9353
 Db 9229 GCGGCTGGACTTGTCCGCTTGGTTTACGCGCTGGCTACAGCGGGGAGACATTTATCA 9288
 Qy 9354 GGGTGTGCGGTCGCGACCCCGCTTATGCTCTTGGCTACTCTCTCTCTCTCTCTCTCTCTCT 9413
 Db 9289 GGGTGTCTCANGCCCGCGCGCTGGTCTGCTGTTTGGCTACTCTCTCTCTCTCTCTCTCTCT 9348
 Qy 9414 TAGGCGCTTTTCT 9473
 Db 9349 TAGGCACTACT 9408
 Qy 9474 ACTGTCCT 9533
 Db 9409 CATTTCT 9468
 Qy 9534 TTTTCT 9593
 Db 9469 CT 9528
 Qy 9594 CT 9653
 Db 9529 CT 9588
 Qy 9654 GGTCTGTGAGCGCAGTACTGAGAGAGTGGTAACTGCTCTCTCTCTCTCTCTCTCTCTCTCT 9711
 Db 9589 GGTCTGTGAGCGCAGTACTGAGAGAGTGGTAACTGCTCTCTCTCTCTCTCTCTCTCTCTCT 9646

RESULT 13
 ABR87286
 ID ABR87286 standard; cdna; 12980 BP.
 XX AC ABR87286;
 XX DT 24-SEP-2002 (first entry)
 XX Hepatitis C virus plasmid p90/HCVFlongpu cdna sequence.
 XX

3295 CGTGGGGGAGATACCGCGCGTGGGTGACATCATCAACGGCTTGGCGCTCTGCGCC 3354
 3366 GACTTGGTGGGAGGCTCTCTTGGCCAGCTGATGCTTACCTCCAGGGGTGGAGTC 3425
 3355 GTAGGGCCAGGAGATGCTTGGCCAGCGACGGAATGGTCTCCAAGGGGTGGAGGT 3414
 3426 TTCTCGCCCCCATCACTGTCTAGCCCGAGCAGACAGTGGCCCTTTTGGGCACCATAGTGG 3485
 3415 TGTGGGCCCATCAAGGGGTAGCCCGAGCAGAGAGAGGCTCTTAGGGTGTATATCA 3474
 3486 TGAGCATGACGGGGCCGCGACAGACAGAAAGAGCTGGGAAATTCAGGTCTCTGTCACAG 3545
 3475 CCAGCTGACTGGCCCGGACAAAACCAAGTGGGGTGGAGTCCAGATCGTGTCAACTG 3534
 3546 TCACTCAGTCTCTCTCGGAACATCATCTCGGGGGTTTTTGGAGCTGTCTACCATGGAG 3605
 3535 CTACCCAAACCTTCTCGCAACGTGATCAATGGGGTATGCTGGACTGTCTACACGGGG 3594
 3606 CTGGCAACAGACTCTGGCGGGCTCACGGGGTCCGGTACGACAGATGTACTCCAGTGTG 3665
 3595 CCGGAACGAGGACCATCGATCACCCAGGGTCTCTCATCCAGATGTATACCAATGTGG 3654
 3666 AGGGGACTTGTAGGGTGGCCAGCCCGCTTGGGACTAAATCTTTGGAGCCGTGCACGT 3725
 3655 ACCAAGACCTTGTGGGCTGGCCGGCTCTCAAGGTTCCTGCTATTGACACCTGCACT 3714
 3726 GTGGAGCGGTGACCTGTACTGTGTCAGCGGACGCTGATGTATCCCGGCTCGAGAC 3785
 3715 GCGGCTCTCGGACCTTTACCTGTGTACAGGACGCGGATGTCAATTCCTGCGCGGG 3774
 3786 GCGGGACAAAACGGGGAGCGCTACTCTCCCGAGACCTCTTCCACCTTTGAAGGGTCT 3845
 3775 GAGTGTATAGAGGGGTAGCTGCTTTCGCGCGGCCAATTTCTTGAAGGCTCT 3834
 3846 CAGGAGCCCGGTGTATGCCCGAGGGGACGCTGTGCGAGTCTTTCGGGGAGCTGTGT 3905
 3835 CCGGGGGTCCGCTGTTGTCGCGCGGGACGCGTGGGCTTATTCAGGGCGCGGTGT 3894
 3906 GCTCTCGGGGTGGCTATAGTCAATAGTTTATCCCGGTGAGACACTCGACATGTC 3965
 3895 GCACCGGTGGGTGGCTAAGGGGTGAGCTTTATCCCTGTGGAGAACCTTAGACAAACA 3954
 3966 CCGGCTCCCGACCTTTAGTGAACAGACACACCTCTGTCGCGGACCTATCAGG 4025
 3955 TGAGATCCCGGTGTTCAGGACACTCTCTCCACAGAGTGGCCGAGCTTCCAGG 4014
 4026 TCGGGTACTTGTGATGCCCGGACTGGCAGTGGAAAGACCAAAAGTTCCTGTGCGATATG 4085
 4015 TGGCCACCTGATGCTCCACCGGCGGTAAGAGCACCAAGTTCGCGGTGGGTAGC 4074
 4086 CTGCTCAGGGGTATAGGTGCTAGTCTTAATCCCTCAGTGGCTGCCACCTGGGGTTG 4145
 4075 CAGCCAGGGGTACAAAGGTGTGGTGTCTCAACCCCTCTGTGTCAGACCTGGGCTTTG 4134
 4146 GGGCGTACTTGTCTAAGGCACATGGCATCAATCCCAATTAGGACTGGAGTCAGACTG 4205
 4135 GTGCTTACATGTCCAAAGGCCATGGGGTGTATCTTATATCAGGACCGGGGTGAGACAA 4194
 4206 TGACGACCGGGGGCCCATCAGCTACTCCACATATGGCAAAATTCCTCGCGATGGGGCT 4265
 4195 TTACCACTGGGACCCCATCAGTACTCCACCTAGGCAAGTTCCTTGGCCGAGCGGGT 4254
 4266 GTGCGGGCGGCTTAGACATCATATGTATGTATGATGCAATGCCGTGGACTTACCA 4325
 4255 GCTCAGGAGGTGCTTATGACATAAATTTGTGACGAGTGGCACTCCACGGATGCCAT 4314
 4326 CCATCTCTGCGTCCGAAAGCTCTTGTATCAAGCAGAGACAGTGGGGTCAGACTAATG 4385
 4315 CATCTTGGGATCGGCATGTCCTTGGACCAAGAGAGACTGCGGGGGGAGACTGGTTG 4374
 4386 TGTGCTACAGCTACGCCCCCTGGGTGAGTGAACACCCCGCCCAACATAGAGGAGG 4445

4375 TGCTCGCCACTGCTACCCCTCGGGCTCGTCACTGTGTCTCCATCTTAACATCGAGGAG 4434
 4446 TGGCCCTTGGGAGGAGGAGATCCCTCTCTATGGGAGGGGATTCCTCCCTGTCTTACA 4505
 4435 TTGCTCTGTCCACCGAGAGATCCCTTTTACGGCAAGGCTATCCCTCTCGAGGTGA 4494
 4506 TCAAGGGAGGAGAGACATCTGATCTTCTGCAATTCAAAGAAAAAGTGTGACGAGCTCGCG 4565
 4495 TCAAGGGGAGAGACATCTCATCTTCTGCACTCAAGAGAGAGTGGCAGAGCTCGCG 4554
 4566 CCGCCCTTGGGGTATGGGCTTGAACCTAGTGGCATATCAAGAGGGTGGAGCTCTCG 4625
 4555 CGAAGCTGTGCTCAATGGGCATCAATGCGTGGCTCTACTACCGCGCTCTTGACGTCTG 4614
 4626 TAATACCAACTCAGGAGACGTAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 4685
 4615 TCATCCCGACACGCGGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4674
 4686 CTGGGCACTTTGACTCCGCTGATCGACTGCAAGCTAGCGGTCACTCAAGTGTGAGACTTCA 4745
 4675 CCGGCGACTTCTGACTCTGTGATAGACTGCAACAGTGTGCTCACTCAGACAGTCAATTC 4734
 4746 GTTATAGACCCCATTTACCATTAACCAACACAGATTTGCTCCCAAGACGCTGTCTCACTG 4805
 4735 GCTTTGACCTTACCTTTACCATTTAGACAAACCGCTCCCGCAGGATGCTGTCTCAGGA 4794
 4806 GCCAGCGCGGGTGGAGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGG 4865
 4795 CTCACCGCGGGGAGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4854
 4866 GTGAGCGAGCTCAGGAATGTTTGAAGTGTAGTGTCTGTGAGTGTCTACGACGAGGGG 4925
 4855 GGGAGCGCCCTCCCGCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4914
 4926 CCGCATGTATGAGTCTACCATCGAGACACCGTCAAGGCTCAGGGGCTATTTCAACA 4985
 4915 GTGCTTGTATGAGTCTACCGCGCGGAGACTACAGTTAGGCTACGAGGCTACATGAACA 4974
 4986 CCGCCGGTGTGCTGTGCGCAGACCATCTTGTAGTGTGGAGGAGTGTTCACGGGG 5045
 4975 CCGCGGGCTTCCCGGTGCGCAGGACCATCTTGAATTTTGGGAGGGCGCTTTTACGGGG 5034
 5046 TCACACATAGATCCCACTTCTTCCCAACAAAGCAATCGGGGAGAAATTTTCGAT 5105
 5035 TCATCATATATAGTCCCACTTCTATCCAGACAAAGAGAGTGGGAGAACTTTTCCT 5094
 5106 ACTTAACAGCTTACAGGCTACAGTGTGCTGCTAGGGCCAAAGCCCGCCCGCTGCTGG 5165
 5095 ACCTGCTAGCGTACCAAGCACCGTGTGCTGCTAGGGCTCAAGCCCTTCCCGCATCGTGG 5154
 5166 AGCTCATGTGGAAGTGTGACTCGACTCAGCCAGCCACACTGCTGGGGCCCAACCTCTCC 5225
 5155 ACCAGATGTGAAGTGTGATTCGCGCTTAAACCCCTTCAATGGGCGCAACCCCTGCG 5214
 5226 TGATCCGCTTGGGCTCTGTACCAAGAGGTCAACCTCACAATCCCGTACGAGAAATACA 5285
 5215 TATACAGACTGGGCGCTGTTCAGATGAAGTCACTCCCTGACCGACCAATCACAATAACA 5274
 5286 TGGCCACTGTATGACCGGACCTTGTAGGTCAATGACAGACATAGGGTCTTGGCAGGGG 5345
 5275 TCATGACATGATGTGGCGGACCTGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5334
 5346 GAGTCTTGGCGGCTGCGCGGTATGCTGGGACCGGGTGTGTTTGCATCATCGGCC 5405
 5335 GCGTCTCGGCTGCTGCGCGGTATGCTGCTTCAACAGGCTGCTGCTGCTGCTGCTGCTGCT 5394
 5406 GCTTGCATTAACCAAGCGAGCGGCTGCTGCGCGGACCAAGAGGCTCTCTATAGGGCTT 5465
 5395 GGATGTCTTGTGCGGGAAGCGGCAATTAATCTGACAGGAGGCTTCTTACAGAGCT 5454
 5466 TTGATGAGTGAAGGAATGTGCTCTAGGGGGGCTCTCATTTGAAGAGGGGACGAGTAG 5525
 5455 TCGATGAGTGAAGAGTGTCTCTCAGCACTTACCGTACATCGAGCAAGGAGTGTGCTCG 5514

[illegible]

Qy	3066	CTTTGACGCGCGTACTCTCGTCAGGCTCACGCTCTACTAGAGTGTGCACCATG	4146	GGCGGTACTGTCTAAGGCACATGGCATCAATCCCAACATTAAGGATGGAGTCAGGACTG	4205
Db	3055	GTITGTCTTAAGTCCCTACTCTCGTCGCGGTTTCAAGGCTTCTCCGATCTGCGCGCTAG	4135	GTGCTTACATGTCCAAGGCCCATGGGTGATCTCTTAATATCAGGACCGGGGTGAGAACA	4194
Qy	3126	CAAGGCATCTCGCGGGGGGAGGTAGCTCAGATGGCGTACTAGCCCTTTGGCAGGTGA	4206	TGACGACCGGGGGCGCCATCAGTACTCACAATATGGCAAAATTCCTCGCGGATGGGGCT	4265
Db	3115	CGCGGAAGATAGCGGAGGTCAATACGTGCAATGGCCATCATCAAGTTAGGGGCGCTTA	4195	TTACCACCTGGCAGCGCCCATCAGTACTCCACCTACGGCAAGTTCTTGGCGACGGCGGT	4254
Qy	3186	CTGGCACTTACATCTATGACCACTCACCCCTATGCGGATGGGATGGCTGTAGTGGCTGC	4266	GTGCGGGGGCGCCTACGACATCATATGTGATGAATCCATGCGGTGGACTCTTACCA	4325
Db	3175	CTGGCACTTATGTATTAACCATCTCACCCCTCTCGAGACTGGGGCGCAACCGGCTGC	4255	GCTCAGAGGTGCTTATGACATAAATTTGTGACGATGCCACTCCACGATGCCACAT	4314
Qy	3246	GGCACTTGGCGGTGCGCTTGGAGCTATCATCTTCACTGTCGATGGAGAGAAAGTCAATTG	4326	CCATCCTTGGCATCGGAAACAGTCTTGTATCAAGACAGACAGCTGGGGTCAGACTAACTG	4385
Db	3235	GAGATCTGGCGGTGGCTGTGGAAACAGTCTCTTCTCCGAAATGGAGACCAAGCTCATCA	4315	CCATCTTGGGCATCGSCACTGCTTGTACCAAGACAGAGACTCGGGGGCGAGACTGTTG	4374
Qy	3306	TTCTGGGAGCGGAGACAGTCTGTTGTGGGACATTTTACACGGACTTCCCGTGTCCGCC	4386	TGCTGGCTACAGTACGCCCCCTGGGTGAGTGACAAACCCCCACCCCAACATAGAGAGG	4445
Db	3295	CGTGGGGGCGAGATACCGCGCGGTGGGTGACATCATCAACGCTTGGCCGTCTCTGCC	4375	TGCTGCCACTGCTACCCCTCCGGGCTCGCTCACTGTGTCCATCTCTAATCATCGAGAGG	4434
Qy	3366	GACTTGTGCGGAGGTCTCTCTTGGCCAGCTGATGGCTATACCTCCAGGGGTGGAGTC	4446	TGSCCCTTGGGAGGAGGCGGAGATCCCTTCTATGGGAGGCGATTTCCCTGTCTTACA	4505
Db	3355	GTAGGGGCCAGGAGATACTCTTGGGCGAGCGGACGAATGGTCTCCAGAGGGGTGGAGT	4435	TTGCTCTGTCCACACCGGAGAGATCCCTTTTACCGCAAGGCTATCCCTCTCGAGTGA	4494
Qy	3426	TTCTCGCCCATCATCTTACGCCAGCAGACAGTGGCTTTTGGGCACCATAGTGG	4506	TCAAGGAGGAGAAAGACATCTCTGCTCATTTCAAGAAAGTGTGACGAGCTCCCG	4565
Db	3415	TGCTGGCGCCCATACAGCGGTAGCGCCAGCAGACGAGAGGCCCTCTAGGGGTGATTAATCA	4495	TCNAGGGGGGAGACATCTCATCTTCTGCTCACTCAAAAGAAAGTGCAGAGAGCTCCCG	4554
Qy	3486	TGAGCATGACGGGGCGGCAAGACAGACAGGCTGGGGAAATTCAGTCTCTGTCCACAG	4566	CGGCCCTTGGGGTATGGGCTTGAATCAGTGGGCTATCTACAGAGGGTGTGAGCTCTCG	4625
Db	3475	CCAGCTGACTGGCGGGAACAAAACCAAGTGGAGGTGAGTCCAGATCGTGTCAACTG	4555	CGAAGCTGTGCTCATTTGGGCATCAATGCGGTGGCTTACTACCGCGGTCTTGAAGTCTCTG	4614
Qy	3546	TCATCTAGTCTTCTCGGAACATPCATCTCGGGGGTGTGTGGAATGTCTTCAATGAG	4626	TAATACCAACTCAGGAGAGAGTGTGCTGCTGCCACCGACCGCTCATGACAGGGGTATA	4685
Db	3535	CTACCCAAACCTTCTGCGCAACGTGCATCAATGGGGTATGCTGAGCTGTCTACACGGG	4615	TCATCCCGACACAGCGCGAGTGTGTGCTGTGCGACCATGTCTCATGATGGGCTTTA	4674
Qy	3606	CTGGCAACAGACTCTGGCGGCTCAGGGGTCCGTCACGAGATGTACTCCAGTGTCTG	4686	CTGGGACTTTGACTCCGCTGATGACTGCAACAGTACGCGTCACTCAAGTTGTAGACTTCA	4745
Db	3595	CCGAAACGAGGACCATCGCATACCCCAAGGTCTGTCTATCCAGATGTATACCAATGG	4675	CCGGGACTTTCGACTCTGTGATAGACTGCAACACGCTGTGTCACTCAGACAGTGCATTTCA	4734
Qy	3666	AGGGGACTTTAGTAGGTGGCCAGCCCGCTCGGACTAAATCTTTGGAGCCCGTGACGT	4746	GTTTAGACCCCAATTTACCATTAACACACAGATTTCTCTCTCAAGAGCTGTCTCAGTA	4805
Db	3655	ACCAAGACTTGTGGGCTGGCCGCTCTCAAGGTTCGCGCTCATTTGACACCCCTGCACCT	4735	GCCTTGACCTTACCTTTACCATTTAGACAAACCAACGCTCCCCCAGGATGTCTCTCAGGA	4794
Qy	3726	GTGAGCGGTGACCTGTACTGTGTACGGGGAACGCTGATGTATCTCCCGCTCGAGAC	4806	GCAGGCGCGGGTCCACCGGTAGGGGAGAGCTGGGCACTTTATAGTATGTTTTCACATG	4865
Db	3715	CGGCTCTCTCGGACTTTTACCTGTGTACAGGCAACGCGGATGTCAITTCGCGCGCGC	4795	CTCAAGCGCGGGCAGAGACTGGCAGGGGAGCCAGGCATCTACAGATTTGTGGCACC	4854
Qy	3786	CGCGGACAAACGGGAGCGCTACTCTCCCGGAGACCTTTTCCACCTTGAAGGGTCTCT	4866	GTGAGCGAGCTCAGGAATGTTTTCAGAGTGTAGTGTCTGTGTGAGTGTCTACGACGAGGG	4925
Db	3775	GAGGTGATAGCGGGTAGCTGTCTTTCGCCCCGGCCCAITTCCTACTTGAAGGGTCTCT	4855	GGGAGCGCCCTCCGGCATGTTTCGACTCTGCTCGCTCTCTGTGAGTGTATGACGCGGCT	4914
Qy	3846	CAGGAGCGCGGTCTATGCCCCAGGGCCACGCTGTGCGAGTCTTCCGGCGAGCTGTGT	4926	CCGCTGTTATGAGTCTACACCATCGGAGACCAACCGTCTCAGGCTCAGGGCGTATTTCAACA	4985
Db	3835	CGGGGGTCTCGTGTGTGCCCCGGGGAACGCGGTGGGCTTATTCAGGGCGCGGTGT	4915	GTCTTGGTATGAGTCTCAGCGCGCGGAGACTACAGTTAGGCTACGAGCGTATATGAACA	4974
Qy	3906	GCTCTCGGGCGGTGAGTCAATGATTTTATCCCGTGTGAGACACTCGACATCTGTCA	4986	CGCGCGGTGTGCTGTGCGCAAGACCATCTTCAGTTTGGGAGGAGTCTTTCACCGGCC	5045
Db	3895	GCACCCGTGGAGTGGCTTAAGCGCGTGCATTTATCTCCCTGTGAGAGACCTAGAGACACCA	4975	CCCCGGGCTTCCCGTGTGCGCAGGACCATCTTGAATTTTGGGAGGGGCTTTTACCGGCC	5034
Qy	3966	CGCGGTCCCAACCTTTAGTGACAAACGACACACCTGTGTGTCGCCCGACCTATCAGG	5046	TCAACACATAGATGCCCACTTCTTTTCCCAAAACAAAGCAATCGGGGAGAAATTTCCCAT	5105
Db	3955	TGAGATCCCCGGTGTGTACCGGACAACTCTCTTCCACGAGAGTGTCCCGCAGAGCTTCCAGG	5035	TCATCTATAGATGCCCACTTCTATCCAGACAAAGCAGAGTGGGGAGAACTTTCTCTT	5094
Qy	4026	TCGGGTACTTGCATGCCCGACTGGCAGTGGAAAGAGCACCAAGTTCTCTGTGCGCATATG	5106	ACTTAAACGCTTACAGGCTACAGTGTGCTTAGGGCCAAAGCCCCCCCCCGTCTCTGGG	5165
Db	4015	TGGCCACCTGCTGCTCCACCGGAGCGGTAGAGACCAAGGTCCCGGCTCGGTACG	5095	ACCTGTGTAGGTACCAAGCCACCGTGTGCGCTAGGGCTCAAGGCCCTTCCCGCATCTGTTGG	5154
Qy	4086	CTGCTCAGGGGTATAAAGTGTAGTGTATTAATCCCTCAGTGGCTGCGACCCCTGGGGTTG	5166	ACGCTGTGTGAGTGTGTTGACTCGACTCAAGCCCCACCTCTGTTGGGCCCCCAGACCTCTCC	5225
Db	4075	CAGCCAGGGCTACAAAGTGTGTGTCTCAACCCCTCTGTGTCTCAACCCCTCTGTGTGCTGCAACGCTGGGGCTT	5155	ACCAGATGTGGAAGTGTGTTGATCCGCTTAAACCCACCTTCCATGGGCGCAACACCCCTGC	5214
			5226	TGTACCGCTTGGGCTCTGTGTTACCAACGAGGTCAACCCCTCACAACATCCCGTGTGACGAAATACA	5285

QY 7374 AACAGCTGGCCATTAAAGTCTTTGGCCAGGCCCCCCCCCAAGCGGCGATTCAGGCTTTTCCA 7433
Db CCGAGCTTGCACCAAAAGTTTGGCAGTCTCTCAACTTCGGCAVTTACGGGCAAAATA 7446
QY 7434 CGGGGGCGGGCTGCCGATTCGGCAGTCAGAGCGCTCTGTATGATGATTTGGCCCTTTCCG 7493
Db CGAACATCCTCTG-----AGCCCGCCCTCTTGCTGCCGCCCGACTCCG 7494
QY 7494 AGACAGGTTCCATCTCTTCCATGCCCCCTCTCAGGGGGAGCTTGGAGATCCAGACCTGG 7553
Db ACGTTGAGTCTATTCTTCCATGCCCCCTCTGAGGGGAGCTTGGGATCCGGATCT-- 7552
QY 7554 AGCCTGAGCAGGTAGAGCCCCCAACCCCCCCCCCAGGGGGGGTGGCAGCTTCCCGGCTCG 7613
Db -----CAGCGAGC 7560
QY 7614 ACTCGGGTCTCGTCTACTTGTCTCCGAGAGAGCAGACTCCGCTCGTGTGTCTCCATGT 7673
Db GGTCTGTCTCGACGGTCACTAGTGGGCGCACAGGGAAGATGTCTGTCTCAATGT 7620
QY 7674 CATACTCTCGACCGGGCTCTAATACTCTTGTAGTCCCGAAGAGGAGAGTTACCGA 7733
Db CTATTCTCGACAGGGGCACTGTCAACCCGCTGGCTGGGGAAGAACAAAACTGCCA 7680
QY 7734 TTAAACCCCTTTGAGCACTCCCTGTTCGATATACAAACAGAGTGTACTGTACCAACAA 7793
Db TCAACGACTGAGCACTCTGTCTACGCCATCACAATCTGGTGTATTCACCACTTCAC 7740
QY 7794 AGAGCGCTCACTAAGGCTTAAAGGTAACTTTTGTATGATGAGGATGGAAGTCTCGACTCT 7853
Db GCAGTGTCTCCAAAGGAGAGAAAGTCACTTTGACAGACTGCAAGTTCTTGACAGCC 7800
QY 7854 ACTACGACTCAGTCTTAAAGGACATTAAAGTACGCGCTCCAAAGTCAACGCAAGCTCC 7913
Db ATTACAGGAGCTGCTCAGAGAGGTCAAGCAGCGGGTCAAAAGTGAAGGCTTAATTGC 7860
QY 7914 TCACCATGAGGAGGCTTGCAGTTAAACCCCACTTCTGCAAGATCTAAATATGGT 7973
Db TATCCGTAGAGGAAGCTTGCAGCTGAGCGGCCCAACATTCAAGTCTTGGCT 7920
QY 7974 TTGGGCTAAGGAGTCCCGAGCTTGTCCGGGAGGCGGTAAACCACTCAAGTCCGTGT 8033
Db ATGGGCAAAAGACGTCGTTGCTGCTGCCAAGAAAGCGGTAGCCCACTCACTCCGCTGT 7980
QY 8034 GGAAGGACTCTCTGAGGACTCAGAAACCAATTCACCAACATTATGGCCAAATAAG 8093
Db GGAAGGACTCTCTGGAAGACGTGTACCAACATAGACACTACATCATGGCCAAAGACG 8040
QY 8094 AGGTGTTCTGCGTGGACCCCAAGGGGGGCAAGAAAGCAGCTCGCTTTATCGTTTACC 8153
Db AGGTGTTCTGCGTTACGCTGAGAGGGGGTCTGTAAGCCAGCTCGTCTCATCGTGTCC 8100
QY 8154 CTGACTCGGCGTCAAGGCTCTGAGAGAGATGCGCTTTATGACATTACAAAACTTC 8213
Db CCAGCTTGGGCGTGGCGGTGCGGAGAGATGCGCTTACGACGTGGTTAGCAAGCTCC 8160
QY 8214 CTCAGCGGTGATGGGGCTTCTTATGATTCAGATTCCCGCTCAGCGGTTAGCT 8273
Db CCTGCGCGGTGATGGGAAGCTCTACGGAATTCAAATCTACACAGGAGCGGGTTGAAT 8220
QY 8274 TTCTTTGAAGCATGGCGGAAAGAGACCTATGGTGTTCGTATGATACCCGAT 8333
Db TCCTGTCGAAGCTGGAAGTCCAGAGACCCCGATGGGTCTCGTATGATACCCGCT 8280
QY 8334 GCTTTGACTCAACCGTCACTGAGAGAGACATCAGGACTGAGGAGTCCATATATCGGGCT 8393
Db GTTTTGACTCCAGTCACTGAGAGGACATCCGTACGAGGAGGAAATTTACCAATGTT 8340
QY 8394 GCTCTTGGCGAGGAGGCCCACTGCCATACCTCGTCACTGAGAGACTTACGTG 8453
Db GTGACTGAGACCCCGGCGGTGGCCATCAAGTCCCTCACTGAGAGCTTTATGTTG 8400
QY 8454 GAGGGCTATGTTCAACAGCAAGGGGCAAACTTGGGGTACAGGCGTTGCGGCGCAGCG 8513

Db 8401 GGGGGCTCTTACCAATTCAGGGGGGAAAACTGGCGCTACCGCAGTGCCTCGCGAGCG 8460
QY 8514 GGGTGTCTACCACTAGCATGGGGAACACCATCACATGCTACGTGAAGCTTTAGCGGCTT 8573
Db GGTACTGACAACTAGCTGTGTAACACCTCTCACTTGTCTACATCAAGGGCCCGGCGAGCT 8520
QY 8574 GTAAAGCTCAGGGATAATTCGGCCCAAAATGCTGTATATGCGCGCATGACTTGTGTGTA 8633
Db GTGAGCGCGCAGGGCTCCAGGACTGCACCATGCTCGTGTGTGGCAGGACTTGTAGTGTGA 8580
QY 8634 TCTCAGAAAGCCAGGGGACCCGAGGAGCAGCGGAACTCTGAGAGCTTTACGAGAGCTTA 8693
Db TCTGTGAAGTGTGGGGTCCAGGAGACCGCGAGCTGTGAGAGCTTTACGAGAGCTTA 8640
QY 8694 TGACAGGTATTTCTGCGCCCTCTGTGTGACCCCCCAGACCGGAGATGATCTGGAGCTGA 8753
Db TGACAGGTACTTCCGCCCCCCCCCGGGGACCCCCCAACCAAGAAATACGACTTGGAGCTTA 8700
QY 8754 TAAACATCTTGTCTCTCAAAATGCTGTGTGGCGCTGGGCCCAACAGGCCCGCGCAGATACT 8813
Db TAAACATGCTCTCTCAAGTGTCACTGCTGCCCAACGACGGCGCTGGAAGAGGGTCTACT 8760
QY 8814 ACTGACCAAGAGACCTTACCCTCCAATCGCCCGGCTGCTCTGGGAAACAGTTTACACT 8873
Db ACTTTACCGCTGACCTTCAACCCCTCGAGAGCGCGTGGGAGACAGCAAGACACA 8820
QY 8874 CCCCTGTCAATTCATGGCTGGGAAACATCATCAGTACGCCCGGACCATATGGGCTGCA 8933
Db CTCAGTCAATTCCTGGCTAGGCAACATAATATGTTTGGCCCCCACACTGTGGGCGAGGA 8880
QY 8934 TGGTCTGTATGACACACTTCTCTCATCTCTCATGCTCAAGACAGCTGGACCAAGACC 8993
Db TGATACTGATGACCACTTCTTTAGGCTCTCATAGCCAGGATCAGCTTTGAACAGGCTC 8940
QY 8994 TCAACTTTGAGATGATCGGAGCGGTACTCTCGTGAATCCCTTGGAGACTTCCAGCTATAA 9053
Db TTAACGTGTGAGATCTACGGAGCTCTACTCTCATAGAACCACTGGATCTACTCTCAATCA 9000
QY 9054 TTGAAGGTGTACATGGCTTGAAGCTTTTCTCTGACACATACACTCCCGCAGACTGA 9113
Db TTAAGAGACTTCCATGGCTCAGCGCATTTTCACTCCACAGTTACTCTTCCAGGTGAATCA 9060
QY 9114 CACGGTGGCTTTCAGCCCTCAGAAACTTTGGGCGCCACCCCTCAGAGGCTGGAAGAGCC 9173
Db ATAGGTGGCGCATGCTCAGAAACTTTGGGCTCCGCGCTTGGAGCTTGGAGTGGAGACAC 9120
QY 9174 GGGCAAGTGTGAGTGTGCTCTCTCTCCGTTGGGGGAGAGCGGCTTTCGGGTG 9233
Db GGGCGCGAGCGTCCCGCTAGGCTTCTGTCCAGAGGAGGAGGCTGTCTATATGTGGCA 9180
QY 9234 GATATCTCTTCAATTTGGGGGTGAAGACCAAGCTCAAACTCACTCACTTGGCGGAGCGC 9293
Db AGTACTCTTCAACTGGGCGTGAAGAACAAAGCTCAAACTCACTCCAATAGCGCGCTG 9240
QY 9294 GCTCTCTGATTTATTCAGCTGTTTCACTCGCGCGCGCGGGGCGGACATTTATCA 9353
Db GCGGCTGTGACTTGTTCGGTGTGTTCAAGCTGGCTTACGCGGGGAGACATTTATCA 9300
QY 9354 GGTGTGCGGTGCCGACCCCGCTTATGCTCTTTGGCTACTCTCTACTTTTGTAGGGG 9413
Db GGTGTGTCTATGCTCCCGCGCGCTGTGTCTGTGTTTGTGCTACTCTCTGCTCGTGCAGGG 9360
QY 9414 TAGGCTTTTCTACTTCCCGCTCGGTAGAGCGGACACATAGTACTACTCATAGCTTA 9473
Db TAGGACTTACTCTCTCCCAACCGATGAAGGTGGGGTAAACACTCCCGGCTCTTAAAGC 9420
QY 9474 ACTGTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9533
Db CAFTTCTGTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9480
QY 9534 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9593

Db	5167	ACCAGATGTGGAAGTGTGTTGATCGCCCTTAAACCCACCCCTCCATGCGGCCAACACCCCTGCG	5226
Qy	5226	TGTACCGCTTGGGCTCTGTTACCAAGAGGTACCCCTCACACATCCCGTGACCAATACA	5285
Db	5227	TATACAGACTGGGCGCTGTTTCAAGATGAAGTCACTGAGCCACCCATCACCANAATACA	5286
Qy	5286	TCGCCACCTGCATGCAAGCCGAGCTTGTAGGTCAATGACACACATGGGTCTTGGCAGGG	5345
Db	5287	TCATGACATGCATGTGCGCCGAGCTGAGGTCTGTCAGGACACCTGGTCTGTTGGCG	5346
Qy	5346	GAGTCTTGGCGGCGCTGCGCGGTATGCTTGGGACCGGGTGTGTTGATCATCGGCC	5405
Db	5347	CGTCTCGGCTGTCTGGCGCGGTATGCTGCTTCAACAGGCTGCGTGTATAGTGGCA	5406
Qy	5406	GCTTGCACATTAACACGAGGCGTCTGTCGCGGACAGGAGTCTCTATGAGGCTT	5465
Db	5407	GGATGCTTGTTCGGGAAGCCGCAATATATCTGACAGGAGGTCTCTACAGAGT	5466
Qy	5466	TTGATGAGATGGAGGAATGTCCCTCTAGGCGGCTCTCATTTGAAGAGGGGACGCGGATAG	5525
Db	5467	TCGATGAGATGGAAGAGTGTCTCAGCACCTTACGGTACATCGAGCAAGGATGATGCTCG	5526
Qy	5526	CCGAGATGCTGAAGTCCAAGATCAAGGCTTATGCAAGAGCTTCCAAACAAGCTCAAG	5585
Db	5527	CTGAGCAGTTCGAAGCAGAGGCGCTCGCGCTCTCTGCAAGCCGCTCCCGCATGCAAGG	5586
Qy	5586	ACATACACCACTGTGCGGCTTCAAGGCCCAAGGTAGAACAACTCTGCGGCCAACAACA	5645
Db	5587	TTATCACCCCTGTCTCAGACCAACTGGCAGAACTCGAGGTCTTTTGGCGGAAGCAACA	5646
Qy	5646	TGTGGAATCTTATAGCGGATCAATACCTCGCAGGACTATCAACACTGCCAGGGAACC	5705
Db	5647	TGTGGAATTTCACTAGTGGGATCAATACTTGGCGGCTGTCAACGCTGCTGTTAACC	5706
Qy	5706	CTGCAATGCTTCAATGATGGGCTTGTAGTGGGCTTCAAGTCCGCTGTCAACAAGCA	5765
Db	5707	CCGCAATGCTTCAATGATGGGCTTGTACAGTGTGCGTCAACAGCCCACTAACCACTGGCC	5766
Qy	5766	CCACTATCTTCTCAACATTTGGGGGCTGGCTAGCATCCCAATTCACCAACCCCGCG	5825
Db	5767	AAACCTCTCTTCAACATATGGGGGGTGGGTGGCTGCCAGCTCGCCGCCCGCGGTG	5826
Qy	5826	GGGCACTGCTCTGTTGTGAGTGGGCTTGTAGTGGGAGCTGCGGTAGGCAGTATAGGCTTAG	5885
Db	5827	CCGCTACTGCTTGTGGTGTGCTGCTAGTGGCGCGCCATCGGCAGCGTTGGACTGG	5886
Qy	5886	GTAAGTGTGTAGTGAATCTGCGAGGGTATGTTGGGGGATTTGGGGGCTCTGCTCG	5945
Db	5887	GGAAGTCTCTGTTGACATTTCTGAGGGTATGCGCGGGCGTGGCGGGAGCTCTTGTAG	5946
Qy	5946	CATTCAAGATCATGCTGCGGAGAGCCCTCCATGGAGGTGCTCAACTTGTGCTGCTG	6005
Db	5947	CATTCAAGATCATGAGCGGTGAGTCCCTCCACGGAGGACCTGGTCAATCTGCTGCCCG	6006
Qy	6006	GAATCTGCTCTCGGGTGCCTTGTGAGTGGAGTCACTGCGCGGCCATTTCTGCGCCGAC	6065
Db	6007	CCATCTCTGCGCTGGAGCCCTTGTAGTGGTGTGCTGCGCAGCAATCTGCGCGCGG	6066
Qy	6066	ACGTGGAGCCGGGGAAGCGCGCTCCAATGGATGAATAGACTCATTTGCTTGTCTTCCA	6125
Db	6067	ACGTTGGCGCGGAGGGGAGTGCATGATGAACCGGCTAATAGCCTTTCGCTCC	6126
Qy	6126	GAGGAATCAGTCCGCGCCACCACTACGTGAGCGAGTCCGATCGCTCGCAGGTGTA	6185
Db	6127	GGGGGAACCAATGTTTCCCGCCACGCACTACGTGCGGAGAGCGATGACAGCGCCCGCTCA	6186
Qy	6186	CCCAACTACTTGGCTCTTACCAATAACACGCTGTGTCAGAAAGTCTCCAACTGGATTA	6245
Db	6187	CTGCCATCTCAGCAGCTCACTGTAAACCGAGCTCTGAGCGACTGCATCAGTGGATAA	6246
Qy	6246	CTGAGGACTGCCCATCCATGCGCGGCTCGTGGCTCCGCGATGTGTGGGACTGGGTTT	6305
Db	6247	GCTCGGAGTGTACCACTCCATGCTCCGGTCTCGGCTAAGGGACATCTGGGACTGGATAT	6306
Qy	6306	GCACCATCTTAACAGACTTTAAAAATTTGGCTGACCTTCCAAATTTATTTCCAAAGATGCCCG	6365
Db	6307	GCGAGTGTGAGGACTTTAGACCTGCTGAAACCAAGCTCATGCCACACTGCTG	6366
Qy	6366	GCTCCCTCTTGTCTCTCTCAAAAGGGGTACAGGGCGTGTGGGCGGACACTGGCATCA	6425
Db	6367	GGATTCCTTGTGTCTGCGAGCGGGGTATAGGGGGGTCTGGCGAGGAGACGGCATTA	6426
Qy	6426	TGACCACACGGTGTCTTGGCGGCCAATATCTCTGCGCAATGTCGCTTGGGCTCCATGA	6485
Db	6427	TGCACACTCGTCCCACTGTGAGCTGAGATCACTGACATGTCAAAACCGGACGATGA	6486
Qy	6486	GAATCAACGGGCGCTAAGACCTTCGATGAATATCTGGCAGGGGACCTTTTCTTATCAATGTT	6545
Db	6487	GGATCGTCTCTAGGACCTGCAGGAACAATGAGGAGTGGGACGTTTCCCATTAACGCTT	6546
Qy	6546	ACAGGAGGCGCAGTGTGCGGAAACCCGCGCAAACTTTAAGTGTGCGCATCTGGAGG	6605
Db	6547	ACACCACGGGCGCTTACTTCCCTTCTGCGCGAACTATTAAGTTGCGCTGTGGAGG	6606
Qy	6606	TGCGCGCTCAGAGTACGCGGAGTGCAGCAGCAGCGGTCAATACCACTACATAACAGGAC	6665
Db	6607	TGTCTCAGAGGAATACGTGGAGTAAGGCGGTGGGGACTTCCACTACGATTCGGGTA	6666
Qy	6666	TCACCACCTGATAACTTGAAGTCCCTGCGCAACTACCTCTCCCGAGTCTTTTCTCTGGG	6725
Db	6667	TGACTACTGACATCTTAATGCGCGTGCAGATCCATCGCCCGAATTTTTCACAGAAT	6726
Qy	6726	TGACGCGAGTGCAGATTCATAGTGTGCGGCCCAACCGAAGCGGTTTTCGCGGATGAGG	6785
Db	6727	TGACGCGGTGCGCTTACACAGGTTTGGCGCCCTTTCGAAGCGCTTGTCTGCGGAGGAGG	6786
Qy	6786	TCTCGTTCTGCGTGGGCTTAATTCATTTGTCTGCGGTCCAGCTTCTCTGCGACCGCTG	6845
Db	6787	TATCATTCAGAGTAGGACTCCAGGATACCGGTGGGTGCGAATTAACCTTTCGAGCGCG	6846
Qy	6846	AACCCGACACAGACGTAATTTGATGCTCATGCTAAACAGATCCATCTCATATCACGCGGAGA	6905
Db	6847	AACCGGAGTAGCGGTGACGCTCATGCTCATGCTTCCATCTCCATATAACAGCAGAGG	6906
Qy	6906	CTGCGCGCGCTTGTAGCGCGGGGTGACCCCGATCCGAGGCAAGCTCTCTGCGGAGCC	6965
Db	6907	CGCGCGGAGAAAGTTGGCGAGAGGTGACCCCTTCTATGGCGAGCTCTCTGCGCTAGCC	6966
Qy	6966	AGCTATCGGACCATCGCTGCGAGCCACTGACACCCACCGCAAGCTATGATGCTGG	7025
Db	6967	AGCTGTCCGCTCCATCTCTCAAGGCACTTTCACCGCAACCAATGACTCCCTCCGACCGG	7026
Qy	7026	ACATGTTGATGCTAACT-----GTTTCATGGGGGCGATGTGACTCGGATAG	7073
Db	7027	AGCTCATAGGCTTAACCTCTGCGGAGGAGATGGCGGCAACATCACCGGGTGG	7086
Qy	7074	AGTCTGGTCCAAAGTGGTGTCTGAGTCTCTGACCCCAATGTTGCGAAGAAAGAGCG	7133
Db	7087	AGTCAGAGAAACAAAGTGTGATTTCTGAGCTCTCTGAGTCTGTCGCGAGAGGAGATG	7146
Qy	7134	ACCTTGAGCTTCGATACCATCAGATACATGCTCCCCCAAGAGAGTCCCAACAGCTT	7193
Db	7147	AGCGGAGGTCTCCGTACCTGCGAGAAATTCGCGGAGTCTCGGAGATTCGCGCGGCC	7206
Qy	7194	TACCGGCTTGGGCAACCGCTGATTACAAACCCAGCTTGTGGAATCTGTGAAAGGCGCAG	7253
Db	7207	TGCGGCTCTGGGCGCGCGGACTCAACCCCGCTGTAGAGAGCTGTGGAAGAGCTG	7266
Qy	7254	ATTACCAACCGGCACTGTTGGGGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG	7313
Db	7267	ACTACGAACCACTGTGTTCCATGGTCCCGCTTACCACTCCACGCTCCCTCTCTGTC	7326
Qy	7314	CTCCCCAAGGAGGCGCGGAGTGGCTTGTGAGGACTCCATAGGAGTCCCTCTTC	7373
Db	7327	CTCGGCTCGGAAAAAGCGTACGGTGTCTCTCCCGAATCAACCTTATCTACTGCTTGG	7386

QY 3006 TGTTGACATACCAAGTGGCTCTTGCGGTGCTTGCGCTGCTTACCTCTTAAAGGTG 3065
DB TAATTTGACATCAACAACTACTCTGGCCATCTTGGACCCCTTTGGAATCTTCAAGCCA 3066
QY 3066 CTTTGACCGCGTGGCTTCTGTCAGGCTCAGCTCTACTGAGAGTGTGACCATGG 3125
DB GTTTGCTTAAAGTCCCTTACTCTGTCGCGCTTCAAGCCCTTCTCCGATCTGCGCTAG 3126
QY 3126 CAAGGCATCTCGCGGGGAGGTAGCTCCAGATGCGCTACTAGCCCTTGGCAGGTGGA 3185
DB CGCGGAAGTAGCGGAGGTCACTAGCGAAATGGCCATCATCAAGTTAGGGGCGCTTA 3186
QY 3186 CTGGCACTTACATCTATGACCACTCAACCTATGTGCGATGGGCTGCTAGTGGCTGC 3245
DB CTGGCACTATGTATTAACCATCTCAACCTCTCTCGAGACTGGCGCAACCGGCTGC 3246
QY 3246 GGGACCTGGCGGTGCGGCTGAGCCATCATCTTCACTCGATGGAGAGAAAGTCAATG 3305
DB GAGATCTGGCGGTGGCTGGAAACCAAGTCTCTTCTCCGAAATGGAGACCAAGCTCATCA 3306
QY 3306 TCTGGGAGCGGAGACAGCTGCTTGTGGGACATTTTACACGACTTCCCGTGTCCGCC 3365
DB CGTGGGGGGAGATACCGCCGCTGCGGTGACATCATCAAGGCTTGCCTCTCTGCC 3366
QY 3366 GACTTGTGCGGAGTCTCTCTTGGCCAGCTGATGCTATACCTCAAGGGGTGGAGTC 3425
DB GTAGGGGCCAGGAGATCTGTTGGCCAGCCGACGGAATGGTCTCCAAGGGGTGGAGT 3426
QY 3426 TTCTGCGCCCATCTGCTTACCGCCAGACAGCTGGCTTGGGACCATAGTGG 3485
DB TGCTGGCGCCATCAGCGCTAGCGCCAGCAGAGAGGCTCTTAGGGGTATTAATCA 3486
QY 3486 TGAGCATAGCGGGCGGACAGACAGACAGCTGGGAAATTCAGGCTCTGTCCACAG 3545
DB CAGCTTGACTGGCCGGGACAAACAAAGTGGAGGTGAGGTCAGATCGTCAACTG 3546
QY 3546 TCACTCAGTCTCTTCGGAACATCCATCTCGGGGTGTTGGAGCTGTCTACCATGGAG 3605
DB CTACCCAAACCTTCTGGCAACGTGCATCAATGGGTATGCTGACTGTCTACCAAGGG 3606
QY 3606 CTGGCAACAGACTCTGCGCGGTCTCAGGGGTCCGTACGAGATGTACTCAGTGTCTG 3665
DB CCGGAACGAGACCATCGCATCAACCAAGGCTCTGTCTATCCAGATGTATACCAATGTG 3666
QY 3666 AGGGGACTTGTAGGTGGCCAGCCCTCGGACTAAATCTTTGGAGCGTGCACCT 3725
DB ACCAAGACTTGTGGGTGGCCCGCTCTCAAGGTTCCCGCTCAITGACACCTGTACCT 3726
QY 3726 GTGGAGCGGTGCACTGTACTCTGTCAACGGAACGTGATGTCTATCCCGGCTCGAAGAC 3785
DB GCGGCTCTCGGACCTTACCTGTCTACGAGGACGCGCATGTCTATCCGTGCGCGGC 3786
QY 3786 GCGGGACAAACGCGGAGCGCTACTCTCCCGAGACCTCTTTCACCTTGAAGGGTCTCT 3845
DB GAGGTGATAGAGGGGTAGCTCTCTTGGCCCGGCGCCATTTCTTCTTGAAGGCTCTCT 3846
QY 3846 CAGAGCGCGGTGTATGCTCCAGGGGACAGCTGTGCGAGTCTTCCGGGAGCTGTGT 3905
DB CCGGGGTCTGCTGTGTGCGCGGACAGCGCTGGGCTTATTCAGGGCGCGGTGT 3906
QY 3906 GCTCTCGGGGCTGGGTCAAGTCTCATAGATTTTATCCCGCTTGAAGACACTCGATCGTCA 3965
DB GCACCGGTGGGTGGCTTAAGCGGTGACTTTATCTCTGTGGAGAACCTTAGGACAAACA 3966
QY 3966 CCGCGTCCCGCATCTTGTAGTGAACAAGACACCACTGCTGTGCCCCAGACCTATCAG 4025
DB TGAGATCCCGGTGTTCAAGCAAACTCTCTCTCCACGAGTGGCCCGAGCTTCCAG 4026
QY 4026 TCGGGTACTTGCATGCCCGACTGGCAGTGGAGAGACCAAGTCTCTGTGCGATATG 4085
DB TGGCCCACTGCTATGCTCCACCGAGCGGTGAAGAGACCAAGGTCCCGGCTGCGTACG 4086
QY 4086 CTGCTCAGGGGTATAAAGTCTAGTGTATTAATCCCTCAGTGGGTGCGCACCCCTGGGTTTG 4145

DB 4087 CAGCCAGGCTTACAAGGTGTTGCTCAACCCCTCTCTGCTGCAACGCTGGGCTTGG 4146
QY 4146 GGGCGTACTTGTCTAAGGCAGATGCAATCAATCCCAAGATTTAGGACTGAGTGCAGACTG 4205
DB GTGCTTACATGTCCAAGGCCCATGGGTTGATCTTAATATCAGGACCGGGGTGAGAACAA 4206
QY 4206 TCAGCAGCGGGCGCCATCAGCTACTCCACATATGCGAAATTCCTTCGCGCATGGGGCT 4265
DB TTACCACTGGCAGCCCATCAGTACTCACCTACGGCAAGTTCCTTGCAGCGCGGT 4266
QY 4266 GTGCGGGCGGCTTACGACATCATATGTGATGAATGCCATGCCGCTGTACCA 4325
DB GCTCAGGAGGTGCTTATGACATAAATTTGTGAOCGAGTGCACCTCCAGGATGCCACAT 4326
QY 4326 CCATCTTGGCATCGGACAGTCTTGTATCAAGCAGACAGCTGGGTCAGACTAATCTG 4385
DB CCATCTTGGGCTTGGGCTGCTTGTACCAAGCAGAGACTCGGGGGCGAGACTGGTTG 4386
QY 4386 TGTGCTACAGTACGCGCCCTGCTGCTCAGTCAACACCCCAACCCCAACATAGAGGAG 4445
DB TGCTCGCCACTGCTACCCCTCGGGCTCGTCACTGTGTCCATCTTAACATCGAGGAG 4446
QY 4446 TGGCCCTTGGGCGAGGCGGAGATCCCTTCTATGGGAGGCGATTCCTCTGTCTTACA 4505
DB TTGCTCTGTCCACACCGGAGAGATCCCTTTTACGGCAAGCTATCCCTCGAGGTGA 4506
QY 4506 TCAAGGAGGAGAGACATCTGATCTTCTGCGCATTCAAAGAAAGTGTAGGACTCGCG 4565
DB TCAAGGGGGGAGACATCTCTTCTGCACTCAAGAAAGTGTGCGAGCTCGCG 4566
QY 4566 CGGCTCTCGGGGTATGGGCTTGAATCAGTGGGCTACTACAGAGGTTGGAGCTCTCG 4625
DB CGAAGTGTGCTGCTTGGGCTCAATCGCTGGGCTACTACCGGCTTGTAGCTGTCTG 4626
QY 4626 TAATACCACTCAGGAGAGCTAGTGGTCTGCCCAACCGACCCCTCATGACAGGGGTATA 4685
DB TCATCCCGACACGCGGATGTTGCTGCTGCGACCGATGCTCTCATGACTGGCTTTA 4686
QY 4686 CTGGGACTTTGACTCTCGGTGATGCTGCAAGCTAGCGGTCACTCAAGTTGTAGACTTCA 4745
DB CCGGCGACTTTCGACTCTGTGATAGACTGCAACCGTGTCTACTCAGACAGTCTGATTTCA 4746
QY 4746 GTTTAGACCCCACTTCAACATAACACACAGATTTGCTCTCAAGACGCTGTCTACGTA 4805
DB GCTTTGACCTTACTTTACATTTGAGACAAACCGCTCCCGAGGATGCTGTCTCAGGA 4806
QY 4806 GCGAGCGCGGGTCTGACGCGGTAGGGGAGACTGCGCATTTATAGGTATGTTTCCACTG 4865
DB CTCAACCGCGGGGAGGACTGCGAGGGGAGCGGACATCTATAGATTTGTGGCACCGG 4866
QY 4866 GTGAGGAGCTCAGGATGTTTGAAGTGTAGTGTCTGTGAGTGTCTAGGAGCGAGGG 4925
DB GGGAGCGCCCTCTCGGCACTGTTTCAAGTCTGCTCTCTGTGAGTGTATGACGCGGGT 4926
QY 4926 CCGCATGTATGACTCAGCTCAGCATCGGAGACCACTGCTGAGGCTCAGGGGTATTTCAACA 4985
DB GTGCTTGTATGAGTCTACGCGCGCGAGACTACAGTTAGGCTACGAGGCTACATGAC 4986
QY 4986 CCGCGGCTTGTGCTGTGCGCAAGACCATCTTGTAGTGTGAGGAGGAGGAGTTTCCCGGC 5045
DB CCGCGGGCTTCCGCTGTGCGGAGGACCATCTTGAATTTTGGGAGGCGCTTTTACGGGC 5046
QY 5046 TCACACATAGATGCCACTCTCTTCCAAACAAAGCAATCGGGGGAATTTCCGAT 5105
DB TCACTCATATAGTCCACTTTTATCCAGACAGAGTGTGGGAGAACTTTCTCTT 5106
QY 5106 ACTTAACGCTTACAGGCTCAGTGTGCTAGGCGCAAGCCCGCCCGCTCTGCG 5165
DB ACTGTAGCTTACAGGACCGGTGCGCTAGGCTTCAAGCCCTCCCGCATCTGCGG 5166
QY 5166 ACGTATGTGGAAGTGTGACTCAGCTCAAGCCCACTCGTGGGCGCCCACTCTCTCC 5225

9294 GCCTCTGATTTATCAGCTGTTTACCGTCCGCGCGCGGGGCGACATTTATCACA 9353
 9241 GCCGGCTGGACTTGTCCGGTTGGTTTACCGCTGGCTACAGCGGGGAGACATTTATCACA 9300
 9354 GCGTGTGCGTGGCGGAGCCCGCTTATGCTCTTTGGCTTACTCTTGTGAGGG 9413
 9301 GCGTGTCTCATGCGCGCGCGCGCTTGTGTTTGGCTTACTCTTGTGAGGG 9360
 9414 TAGGCTTTTCTTACTCCCGCTCGGTAGAGCGGCACATAGCTACATCATAGCTA 9473
 9361 TAGGCTACTACCTCTCTCCCAACCGATGAAGTTGGGTAAACACTCCGCGCTCTTAAGC 9420
 9474 ACTGTCCTT 9533
 9421 CATTTCTGTTC 9480
 9534 TT 9593
 9481 TTCTTTTTTCTTCTCT----- 9497
 9594 CTTATTTTACTTTCTTCTGTTGGTCCATCTTAGCCCTAGTCACGGCTAGCTGAGAA 9653
 9498 -----TTTTCCTTCTTTTAAATGGTGGCTCCATCTTAGCCCTAGTCACGGCTAGCTGAGAA 9553
 9654 GGTCCGTGAGCCGCTGACTGACAGAGTSCCGTAACTGTCTCTCTGAGATCATGT 9711
 9554 GGTCCGTGAGCCGCTGACTGACAGAGTGTCTGATCTGCTCTCTGAGATCATGT 9611

RESULT 10
 AAC86648
 ID AAC86648 standard; DNA; 9611 BP.
 AC AAC86648;
 XX
 XX
 02-APR-2001 (first entry)
 DE Nucleotide sequence of chimeric Hepatitis C virus clone pH77 (p7)-J65.
 XX HCV; HCV strain HC-J6CH; HCV genotype 2a; antiviral; vaccine; ss.
 XX Synthetic.
 OS Hepatitis C virus.
 XX
 XX
 Key Location/Qualifiers
 CDS 342..9389
 /*tag= a
 WO200075338-A2.
 14-DEC-2000.
 XX
 XX
 02-JUN-2000; 2000WO-US15446.
 XX
 04-JUN-1999; 99US-0137693.
 XX
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES;
 XX Yanagi M, Bukh J, Emerson SU, Purcell RH;
 WPI: 2001-061728/07.
 P-PSDB; AAB30733.
 XX
 Nucleic acid molecule encoding human hepatitis C virus of genotype 2a
 PT for developing vaccines, for diagnosis of hepatitis C virus and in
 PT screening assays for identification of antiviral agents -
 XX
 PS Disclosure; Page 146-149; 167pp; English.
 XX
 AAC86645-48 represent chimeric cDNA clones of infectious Hepatitis C
 CC virus (HCV). In each clone the C, E1 and E2 genes are derived from HCV
 CC strain HC-J6CH genotype 2a. The p7 protein was derived from HCV strain

CC HC-J6CH or HCV strain pCV-H77C, and the NS genes were derived from HCV
 CC strain pCV-H77C. Such HCV sequences are capable of expressing the virus
 CC when transfected into cells. The HCV protein is useful for assaying
 CC candidate antiviral agents for activity against HCV. Antibodies specific
 CC for HCV polypeptide are useful in prevention and treatment of diseases
 CC caused by HCV in animals, in particular humans. The HCV polypeptides
 CC serve as immunogens in the development of vaccines for preventing HCV
 CC in mammals or as antigens in diagnostic assays for detecting the
 CC presence of HCV in biological samples. The HCV polynucleotide is also
 CC useful for identifying cell lines capable of supporting the replication
 CC of HCV in vitro and to produce attenuated viral strains via passage in
 CC vitro or in vivo.
 XX
 SQ Sequence 9611 BP; 1925 A; 2897 C; 2695 G; 2094 T; 0 other;
 Query Match 58.4%; Score 5676; DB 22; Length 9611;
 Best Local Similarity 74.6%; Pred. No. 0;
 Matches 7248; Conservative 0; Mismatches 2345; Indels 125; Gaps 4;
 QY 6 CCCCTAATAGGGGCGACACTCCGCCATGAATCACTCCCTGTGAGGACTACTGTCTCA 65
 DB 7 CCCCTGATGGGGGCGACACTCCCATGAATCACTCCCTGTGAGGAACTACTGTCTCA 66
 QY 66 CGCAGAAAGCGTCTAGCCATGCGTGTAGTATGAGTGTGTACAGCTCCAGGCCCCCC 125
 DB 67 CSCAGAAAGCGTCTAGCCATGCGTGTAGTATGAGTGTGTGAGCTCCAGGACCCCC 126
 QY 126 TCCCGGAGAGCCATAGTGTCTCGGAAACCGGTGAGTACACCGGAATTCGCGGAGAC 185
 DB 127 TCCCGGAGAGCCATAGTGTCTCGGAAACCGGTGAGTACACCGGAATTCGCGGAGAC 186
 QY 186 TGGGTCTTTCTTGGATAAACCCACTCTATGCCCGGCCATTTGGGCGTCCCGCAAGA 245
 DB 187 CGGGTCTTTCTTGGATAAACCCCGTCAATGCTGAGATTTGGGCGTCCCGCAAGA 246
 QY 246 CTGCTAGCCGAGTAGCTGTGGTGTGCGAAAGCCCTTGTGGTACTGCTGATAGGTGCTT 305
 DB 247 CTGCTAGCCGAGTAGCTGTGGTGTGCGAAAGCCCTTGTGGTACTGCTGATAGGTGCTT 306
 QY 306 GCGAGTCCCGGAGGTCTGTAGACCGTGCACCATGAGCAAAATCTTAAACCTCAAA 365
 DB 307 GCGAGTCCCGGAGGTCTGTAGACCGTGCACCATGAGCAAAATCTTAAACCTCAAA 366
 QY 366 GAAACCAAGAAAGAACCAACCCCTGCCCAACAGAGCTTAAGTTTCGGGCGCGGCC 425
 DB 367 GAAACCAAGAAAGAACCAACCCCTGCCCAACAGAGCTTAAGTTTCGGGCGCGGCC 426
 QY 426 AGATCGTTGGCGAGTATATCTTTGTTGCCGCGAGGGGCCCCAGGTGTGGTGTGCGCGA 485
 DB 427 AGATCGTTGGCGAGTATATCTTTGTTGCCGCGAGGGGCCCCAGGTGTGGTGTGCGCGA 486
 QY 486 CAAGGAAGACTTTCGGAGCGGTCCAGCCACGTTGGAAGGGCGCCAGCCCATCCCTAAAGATC 545
 DB 487 CAAGGAAGACTTTCGGAGCGGTCCAGCCACGTTGGAAGGGCGCCAGCCCATCCCTAAAGATC 546
 QY 546 GCGCGTCCACTGGCAAAATCTTGGGAAACCAAGGATACCCCTGGGCCCTATACGGGAATG 605
 DB 547 GCGCGTCCACTGGCAAAATCTTGGGAAACCAAGGATACCCCTGGGCCCTATACGGGAATG 606
 QY 606 AGGGAAGTCCGCTGGCGAGGATGCTCTCTCTCCCGGAGGTTCCCGTCTCTTGGGGCC 665
 DB 607 AGGGAAGTCCGCTGGCGAGGATGCTCTCTCTCCCGGAGGTTCCCGTCTCTTGGGGCC 666
 QY 666 CCAATGACCCCGCATAGGTTCGCGCAACGTTGGGTAAAGTCAATCATACCTTAAAGTGG 725
 DB 667 CCAATGACCCCGCATAGGTTCGCGCAACGTTGGGTAAAGTCAATCATACCTTAAAGTGG 726
 QY 726 GCTTTGCGGACCTCATGCGGTACATCTCTGTGCGGCGCCCGCTCGCGCGCTCGCA 785
 DB 727 GCTTTGCGGACCTCATGCGGTACATCTCTGTGCGGCGCCCGCTCGCGCGCTCGCA 786
 QY 786 GAGCTCTTCGCGCATCGGTGAGGATCTCTGAGGACGGGTGTTAAATTTTGCACAGGAACT 845

6066	QA	GAATCTGTCTCCGGTGCCTTGGTAGTGGAGTCAATCTGCGCGGCACATCTTGCGCGCAG	6066
6007	DB	CCATCTCTCGCTCGAGCCCTTGTAGTGGTGTGTCTGCGCAGCAATACTCTGCGCGCGC	6066
6066	QA	ACGTGGGACCGGGGAAGCGCGCTCCAAATGATGAATAGACTCAATGCTTTGCTTCCA	6125
6067	DB	ACGTTTGGCCCGGCGAGGGGCAGTGCAATGATGAACCGGCTAATAGCTTTCGCTCCC	6126
6126	QA	GAGGAATACACGTCCGCCCCACCCACTACTAGTGAACGAGTTCGGATCGCATCGTGTGA	6185
6127	DB	GGGGGAAACCATGTTTCCCCACGCACTAGCTGCGGAGAGCGATGAGCGCCCGCTCA	6186
6186	QA	CCCAACTACTTGGCTCCCTTACCAATACACAGCTCTCAGAGAATCCAAACTGGATTA	6245
6187	DB	CTGCCATACTCAGCAGCCCTCACTGTAAACCCAGCTCTGAGGCGACTGCATCAGTGGATAA	6246
6246	QA	CTGAGGACTGCCCATCCGATCCGATGCGCGGCTCGTGGCTCGCGCATGTGCGGACTGGTTT	6305
6247	DB	GCTCGGAGTGTACCACTCCATGCTCCGGTCTCTGGCTAAGGAGCATCTGGGACTGGATAT	6306
6306	QA	GCACCATCTTAACAGACTTTAAAAATTCGCTGACCTCCAAATTAATCCCAAGATGCCCG	6365
6307	DB	GGAGGTGCTGAGCGACTTTAAGACTGGCTGAAGCCAGCTCATGCCAACTCTGCTG	6366
6366	QA	GCCTCCCTTTGTCTCTGTCAAAAGGGGTACAAGGGCTGTGGGCCCGCACTTGCATCA	6425
6367	DB	GGATTCCTTTGTCTCTCCAGCGCGGTATAGGGGGTCTGGCGAGGAGCGCATTTA	6426
6426	QA	TCACACAGGTTGCTCTTGGCGGCCAATATCTCTGGCAATGTTCGCTTGGGCTCGATGA	6485
6427	DB	TGCACACTCGTCCCACTCTGGAGCTGAGTCACTTGGACATGTCAAAAACGGGACGATGA	6486
6486	QA	GAATCAACGGGGCTTAAGACTTCATGAATATCTGACAGGGGACCTTCTCTATCAATTTGT	6545
6487	DB	GGATCGTTCGTCTTAGGACCTTCAGAGAACATGTGGAGTGGGACGTTTCCCATTAACGCT	6546
6546	QA	ACACGGAGGGCCAGTGCCTGCGGAAACCGCGGCCAAACTTTAAGTTCGCATCTCGAGGG	6605
6547	DB	ACACCACGGGCCCTGTACTCCCTTCTGCGCGCAACTATAAGTTTCGGCTGTGGAGGG	6606
6606	QA	TGGCGGCTCTAGAGTACCGGAGGTGACGACGACCGGTTCATACCACTACATATAACAGAC	6665
6607	DB	TGCTTCAGAGGAATACCTGGAGATAAGCGGGTGGGGGACTTCCACTACGTATCGGGTA	6666
6666	QA	TCACCACTGATTAACCTTGAAGTCCCTGCCAACTACCTCTCCAGTTCCTTTCTCTGG	6725
6667	DB	TGACTACTGACATCTTAAATGCCCGTGCAGATCCCATCGCCGAAATTTTTCACAGAAAT	6726
6726	QA	TGGACGGAGTGACATCCATAGTTTGGCCCCCACACCGAAGCGGTTTTTTCGGGATGAGG	6785
6727	DB	TGGACGGGTGGCCCTACACAGTTTGGCCCCCTTGCAGCCCTTGTCTGGGAGGAGG	6786
6786	QA	TCTGTTCTGGTGGGCTTAATCATTTGTGTGGGTTCAGCTTCTTGTGCAACCTTG	6845
6787	DB	TATCATTCAGAGTAGACTCCACAGTATACCGGTGGGGTCGCAATTAAGCTTCGAGCCCG	6846
6846	QA	AACCGGACACAGCGTATTGATGCCATGCTAAACAGATCCATCTCATATACGCGGAGA	6905
6847	DB	AACCGGAGTACCGGTGTGACGTCCATGCTCACTGATCCCTCCCATATAACAGCAGAGG	6906
6906	QA	CTGACGCGCGGCTTTAGCGGGGGTCAACCCCATCCGAGCAAGCTCTTCGGCGAGCC	6965
6907	DB	CGGCGGAGAGGTTGGCGAGAGGTTCAACCCCTTCTATGGCCAGCTCTTCGGCTAGCC	6966
6966	QA	AGCTATCGGCACCATCGCTGCGAGCCACTGCACCCACCGCAGGCAAGCCTTATGATGG	7025
6967	DB	AGTGTCCGCTCCATCTCTCAAGGAACTTGCACGCCCAACCACTGACTCCCTGACGCG	7026
7026	QA	ACATGGTGGATGCTAACTT-----GTTTCATGGGGGGGATGTGACTCGGATAG	7073
7027	DB	AGCTCATAGAGGCTAAACCTCTGTGGAGGCGAGGATGGGCGGCAACATACCAAGGGTTG	7086
7074	QA	AGTCTGGGTCCAAAGTGGTCTCTTCTGACTCTCTCGACCAATGGTCCGAAGAAAGGAGCG	7133

8994 TCAACTTTGAGATGTACGAGCGGTGTACTCCGTGAGTCCCTTGGACCTCCACGCTATAA 9053
 8941 TTTAACTGTGAGATCTACGAGCGCTGTACTCCATAGAACCACTGGATCTACTCCCAATCA 9000
 9054 TTGAAAGTTTACATGGCTTGAAGCTTTTCTCTGTGACACATACACTCCGCCAGAACTGA 9113
 9001 TTCAAGAGCTTCCATGGCTTGAAGCTTTTCTCTGTGACACATACACTCCGCCAGAACTGA 9060
 9114 CAGGGTGGCTTCAGCCCTCAGAAACTTGGGGCGCCACCCCTCAGAGCGTGAAGAGCC 9173
 9061 ATAGGTGGCGGATGCTCAGAAACTTGGGGTCCGCCCTTGGAGCTTGGAGACACC 9120
 9174 GGGCAGCTGTCAGTCAAGGCGCTCCCTCATCTCCGTGGGGGAGAGCGCGTTTGGCGTC 9233
 9121 GGGCGGAGCGCTTGGCTTGGCTTCTGCTCAGAGAGGAGGAGGCTGTATATGTGGCA 9180
 9234 GATATCTTCAATTTGGGGCGGTGAAGACCAAGCTCAAACTCACTCCATTCGCCGAGCGC 9293
 9181 AGTACCTCTTCAACTGGGCGAGTAAGAACAAAGCTCAAACTCACTCCCAATAGCGCGCTG 9240
 9294 GCCTCTGGATTTATCCAGCTGGTTTCACTCGTGGCGCGGGGGGAGACATTTATCACA 9353
 9241 GCGGCTGGAGCTTGTCCGGTTGGTTTCAAGCTGCTACAGCGGGGAGACATTTATCACA 9300
 9354 GCGTGTCCGCTGCGGACCCCGCTTATTGCTCTTTGGCTTACTCTCTTTTGTAGGGG 9413
 9301 GCGTGTCTCATGCGCGCGCGCTGTCTGTGTTTGGCTTACTCTCTCTCTCTCTCTCT 9360
 9414 TAGGCTTTTCTACTCTCCGCTCGGTGAGAGCGGACACATTTAGCTTACATCTCCATAGCTA 9473
 9361 TAGGCATCTACTCTCTCCCAACCGAGTGAAGTTGGGGTAAACACTCCGCGCTCTTTAAGC 9420
 9474 ACTGTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9533
 9421 CATTTCTGTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTT 9480
 9534 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 9593
 9481 TTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 9497
 9594 CTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 9653
 9498 ----TTTCTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9553
 9654 GGTCCGTGAGCGCATGCTGAGAGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 9711
 9554 GGTCCGTGAGCGCATGCTGAGAGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 9611

RESULT 9
 AAC86646
 ID AAC86646 standard; DNA; 9611 BP.
 AC AAC86646;
 XX
 XX
 XX 02-APR-2001 (first entry)
 DT
 XX
 DE Nucleotide sequence of chimeric Hepatitis C virus clone pH77(p7)-CV-J6S.
 XX HCV; HCV strain HC-J6CH; HCV genotype 2a; antiviral; vaccine; ss.
 XX
 OS Synthetic.
 OS Hepatitis C virus.
 XX
 FH Key Location/Qualifiers
 FT CDS 342..9389
 FT /*tag= a
 XX
 XX WO200075338-A2.
 XX
 XX 14-DEC-2000.
 XX
 XX 02-JUN-2000; 2000WO-US15446.
 PF

XX
 PR
 XX
 PA 04-JUN-1999; 99US-0137693.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Yanagi M, Bukh J, Emerson SU, Purcell RH;
 XX WPI; 2001-061728/07.
 XX P-PSDB; AAB30731.
 XX
 XX Nucleic acid molecule encoding human hepatitis C virus of genotype 2a
 PT for developing vaccines, for diagnosis of hepatitis C virus and in
 PT screening assays for identification of antiviral agents -
 XX
 XX Disclosure; Page 115-118; 167pp; English.
 XX
 XX AAC86645-48 represent chimeric cDNA clones of infectious Hepatitis C
 CC virus (HCV). In each clone the C, E1 and E2 genes are derived from HCV
 CC strain HC-J6CH genotype 2a. The p7 protein was derived from HCV strain
 CC HC-J6CH or HCV strain pCV-H77C, and the NS genes were derived from HCV
 CC strain pCV-H77C. Such HCV sequences are capable of expressing the virus
 CC when transfected into cells. The HCV protein is useful for assaying
 CC candidate antiviral agents for activity against HCV. Antibodies specific
 CC for HCV polypeptide are useful in prevention and treatment of diseases
 CC caused by HCV in animals, in particular humans. The HCV polypeptides
 CC serve as immunogens in the development of vaccines for preventing HCV
 CC in mammals or as antigens in diagnostic assays for detecting the
 CC presence of HCV in biological samples. The HCV polynucleotide is also
 CC useful for identifying cell lines capable of supporting the replication
 CC of HCV in vitro and to produce attenuated viral strains via passage in
 CC vitro or in vivo.
 XX
 SQ Sequence 9611 BP; 1924 A; 2898 C; 2694 G; 2095 T; 0 other;
 Query Match 58.6%; Score 5692; DB 22; Length 9611;
 Best Local Similarity 74.7%; Pred. No. 0;
 Matches 7258; Conservative 0; Mismatches 2335; Indels 125; Gaps 4;
 QY 6 CCCTTAATAGGGCGGACACTCCGCCATGAATCACTCCCTGTGAGAACTACTGTCTTCA 65
 DB 7 CCCTGATGGGGCGGACACTCCACCATGAATCACTCCCTGTGAGAACTACTGTCTTCA 66
 QY 66 CGCAGAAAGCGTCTAGCCATGGCGTTGATGATGATGATGATGATGATGATGATGATG 125
 DB 67 CGCAGAAAGCGTCTAGCCATGGCGTTGATGATGATGATGATGATGATGATGATGATG 126
 QY 126 TCCCGGAGAGCCTATGCTGTGCGGACCGGTGAGTACACCGGAATTCGCCGGAGAGAC 185
 DB 127 TCCCGGAGAGCCTATGCTGTGCGGACCGGTGAGTACACCGGAATTCGCCGGAGAGAC 186
 QY 186 TGGGTCTCTTTTGGATAAAGCCACTCTATGCGCGGCCATTTTGGGGGTGCGCCCGCAAGA 245
 DB 187 TGGGTCTCTTTTGGATAAAGCCACTCTATGCGCGGCCATTTTGGGGGTGCGCCCGCAAGA 246
 QY 246 CTGCTAGCCGAGTAGCGTTGGGTGCGAAAGCGCTTGTGATGATGATGATGATGATGATG 305
 DB 247 CTGCTAGCCGAGTAGCGTTGGGTGCGAAAGCGCTTGTGATGATGATGATGATGATGATG 306
 QY 306 GCGAGTGCCTCCCGGAGGCTCTGCTAGACCGGTGACCATGAGCACAATCTTAAACCTCAA 365
 DB 307 GCGAGTGCCTCCCGGAGGCTCTGCTAGACCGGTGACCATGAGCACAATCTTAAACCTCAA 366
 QY 366 GAAAAACCAAAAGAAACACCAACCGTCCGCCCAAGAGCTTTAAAGTTTCCGGGGCGGCGCC 425
 DB 367 GAAAAACCAAAAGAAACACCAACCGTCCGCCCAAGAGCTTTAAAGTTTCCGGGGCGGCGCC 426
 QY 426 AGATCGTGTGCGAGTATATCTTGTTCGCGCGCAGGGGCCCGCAGGTTGGGTGTGGCGCGA 485
 DB 427 AGATCGTGTGCGAGTATATCTTGTTCGCGCGCAGGGGCCCGCAGGTTGGGTGTGGCGCGA 486
 QY 486 CAAGGAAGACTTCGGAGCGGTCCAGCCAGCTGGAGGGCCCGCCATCCCTAAAGATC 545
 DB 487 CAAGGAAGACTTCGGAGCGGTCCAGCCAGCTGGAGGGCCCGCCATCCCTAAAGATC 546

Copied from 09540843 on 05/19/2004

QY 4626 TAATACCACTCAGGAGAGCTAGTGTCTGCTGCGCACCGAGCCCTCATGACAGGGTATA 4685
 DB 4627 TCATCCCGACCGGCGATGTTGTCGTGTCGACCGATGCTCTCATGACTGGCTTTA 4686
 QY 4686 CTGGGAGCTTTGACTCGGTGATCGACTGCAAGCTAGGCGTCACTCAAGTTGTAGACTTCA 4745
 DB 4687 CCGGCACTTCGACTCTGTGATAGACTGCAACACGCTGTCTCACTCAGACAGTGCATTTCA 4746
 QY 4746 GTTTAGACCCACATTCACCATTAACACACAGATGTCCTCAAGAGCTGTCTCAGCTA 4805
 DB 4747 GCCTTGACCTCACTTTTACCATTTGAGACAAACAGCTTCCCGAGAGTCTGTCTCAGGA 4806
 QY 4806 GCCAGCCCGGGGTGCGACGGGTAGGGAAGACTTGGGCATTTATAGGTATGTTTCCACTG 4865
 DB 4807 CTCAAACCGCGGAGGAGCTGGCAGGGGGAAGCCAGGCATCTATAGATTTGTGCGACCCG 4866
 QY 4866 GTGAGCAGCTCAGGAATGTTTGAAGTGTAGTCTCTGTGAGTGTCTGAGTGTCTGAGCGAGGGG 4925
 DB 4867 GGGAGCGCCCTCCGGCATGTTGACCTCGCTCCGTCTGTGAGTGTATGACCGCGGGCT 4926
 QY 4926 CCGCATGTATGAGTCCACACCATCGGAGACACCGTCAAGCTCAGGGCTGTTTCAACA 4985
 DB 4927 GTGCTTGTATGAGCTCACCGCCCGCAGACTACAGTTAGGCTACGAGCGTATGNAACA 4986
 QY 4986 CGCCCGGTTTGCCTGTGTGCAAGACCATCTTGAGTTTGGGAGGCAAGTTTTCACCGGCC 5045
 DB 4987 CCCCGGGCTTCCGCTGTGCGAGGACCATCTTGAAATTTGGGAGGGGCTCTTTACGGGCC 5046
 QY 5046 TCACACACATAGATGCCACTTCTTCCCAAACAAAGCAATCGGGGGGAAATTTTCGCAT 5105
 DB 5047 TCACTCATATAGATGCCACTTTTATCCAGACAAAGCAGAGTGGGGGGAACCTTTCCCTT 5106
 QY 5106 ACTTAACAGCTACAGGCTACAGTGTGCGCTAGGCGCAAGCCCCCCCCCTGCTGGG 5165
 DB 5107 ACTGTGTAGGTACCAAGACCACTGTGCGCTAGGGCTCAAGCCCCCTCCCTCATCTGTGGG 5166
 QY 5166 ACCTGTGTGGAGTGTGACTCGACTCAAGCCCAACACTCGTGGGCCCCACACCTCTCC 5225
 DB 5167 ACCAGATGTGGAGTGTGATCCGCTTAAACCCACCTCCATGGGSCAACACCCCTGC 5226
 QY 5226 TGTAACCGTGTGGCTGTGTTACCAACAGAGTCACTTACACATCTCCGCTGACGAAATACA 5285
 DB 5227 TATACAGACTGGGCGCTGTTCAAGATGAAGTCACTGAGCGACCCCAATCACCANAATACA 5286
 QY 5286 TCGCCACTGCATGCAAGCGGCTTCAAGTCAAGCAGCAGCATGAGTGTCTTGGCAGGGG 5345
 DB 5287 TCATGATGATGATCGCGGACCTGAGGCTGTCAGAGACCTGAGGCTGCTGTTGGCG 5346
 QY 5346 GAGTCTTTGGGCGCGTGTGCTGCTGCGGACCGGGTGTGTTGATCATCTCGGCC 5405
 DB 5347 GCGTCTGCTGCTGCGCGGCTGTTGCTGCTCAACAGGCTGCTGATAGTGGCA 5406
 QY 5406 GCTTGCATTAACAGGAGCGCTGTGTGGCGCGGACAGAGGCTCTCTATAGAGCTT 5465
 DB 5407 GGATCGTCTTGTCCGGGAAGCGGCAATATATACCTGACAGGAGGTTCTCTTACCAGAGT 5466
 QY 5466 TTGATGAGATGAGGAATGTGCTCTAGGGGCTCTCAATGAGAGGGGCGAGGATAG 5525
 DB 5467 TCGATGATGGAAGAGTGTCTCAGCACTTACCGTATCGACAAAGGAGTATGCTCG 5526
 QY 5526 CCGAGATGCTGAAGTCAAGATCAAGGCTTATGCGAGCAAGCTTCAAAACAAAGCTCAAG 5585
 DB 5527 CTGAGCAGTTCAAGCAGAGGCGCTTCGGCTCTGCGAGCCGCTCCCGCCATGACAGG 5586
 QY 5586 ACATPACCACTGTGAGGCTTCTAGGCGGCTTCAAGAGGTAGAACAAATCTGGGCGCAACACA 5645
 DB 5587 TTATCACCCCTGCTGTCAGACCAACTGGCAGAAACTCGAGGCTCTTTGGGCGGAAGCACA 5646
 QY 5646 TGTGGAATTCATTTAGCGGATCAATACCTCGCAGGACTATCAACTGCCAGGGAACC 5705
 DB 5647 TGTGGAATTTTCATCAGTGGGATACAAATCTTGGCGGGGCTGTCAACGCTGCTGTTAACC 5706
 QY 5706 CTGCAAGTAGCTTCCATGATGCGTTCAGTGGCGGCCCTCACAGTCCGCTGTCAACAAGCA 5765

DB 5707 CCGCCATGCTTCAATGATGGCTTTTACAGTGTGCGCTCACAGCCCACTAAACCACTGGCC 5766
 QY 5766 CCACTATCTTCTCAACATTTTGGGGGCTGCTAGCTAGCTTCCAAATTTGACCAACCGCGG 5825
 DB 5767 AAACCCCTCTCTTCAACATATTTGGGGGGTGGGTGCTGCCAGCTGCGCGCCCCGGTG 5826
 QY 5826 GGGCCACTGGCTTCTGTTGTGCTAGTGGGAGCTGCCGTAGGAGCTATAGGCTTAG 5885
 DB 5827 CCGCTACTGCTTGTGGGTGCTGGCTAGCTGGCGCGCCCATCGSCAGCTTGGACTG 5886
 QY 5886 GTAAGTGTAGTGGACATCTGCGAGGGTATGTTGGCGGCAATTTGCGGGGCTCTCTGTCG 5945
 DB 5887 GGAAGTCTCTGTTGGACATTTCTGAGGGTATGCGCGCGGCTGGCGGGAGCTCTTGTAG 5946
 QY 5946 CATTAAGATCATGTCTGCGGAGAGCCCTTCCATGGAGATGTCGTCAACTTGTCTGCCCTG 6005
 DB 5947 CATTAAGATCATGACGGGTGAGTCCCTTCCAGGAGACCTGCTCAATCTGCTGCCCG 6006
 QY 6006 GAAATCTGTCTCGGGTGCCTTGTAGTGGAGTCACTGCGCGGCTCATTTCTGCGCGGAC 6065
 DB 6007 CCATCTCTCGCTGGAGCCCTTGTAGTGGGTGCTGCTGCGCAGCAATCTGCGCGCGC 6066
 QY 6066 AGTGGGACCGGGGGAAGGCGGCTCCATGTGATGAAATAGACTCATTTGCTTGTCTCCA 6125
 DB 6067 AGTGTGGCGCGGAGGCGGAGTGAATGGAACCGGCTTAATAGCTTTCGCTCC 6126
 QY 6126 GAGGAATCAGCTGCGCCCCCCTAGCTAGTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 6185
 DB 6127 GGGGGAACCATGTTTCCCGCAGCACTAGTGGCGGAGGAGTGGAGCGCGCGCTCA 6186
 QY 6186 CCAACTACTTGGCTCCCTTACCAATAACAGGCTGTCTGAGAGACTTCCAACTGGAATTA 6245
 DB 6187 CTGCCATCTCAGCAGCTCTCATGTAAACAGCTCTTGGAGGAGTGCATCAGTGGATAA 6246
 QY 6246 CTGAGACTGCCCCATCCATCGCGGGCTGTGGTCTCGCGGATGTTGGGAGCTGGGTTT 6305
 DB 6247 GCTCGGAGTGTACCACTCCATGCTCCGTTCTTGGCTTAAGGAGCATCTGGGAGTGGAT 6306
 QY 6306 GCACCATCTACAGACTTTTAAATTTGGTGTGCTCCAAATTTATTTCCAAAGATGCCCG 6365
 DB 6307 GCGAGTGTCTGAGGAGCTTTTAAAGCTTGAAGCTGAAAGCAGCTCATGCCAAGTGGCTG 6366
 QY 6366 GCCTCCCTCTTGTCTCTCTGCAAAAGGGGTACAAGGGCGTGTGGGCGGCGACTTGGCATCA 6425
 DB 6367 GGATTCCTTGTGCTCTGCGCAGCGGGTATAGGGGGTCTGCGGAGGAGACGGCATTA 6426
 QY 6426 TGACCAACAGGCTCTTGGGCGCCATATCTCTGGCAATGTCCGCTTGGGCTCCATGA 6485
 DB 6427 TGCACACTCGCTGCCACTGTGGAGCTGAGATCACTGGACATGTCAAAAACGGGACGATGA 6486
 QY 6486 GATACAGGGGCTTAAAGCTTGAAGTCTGAGCAATATCTGGCAGGGGACTTTTCTTATCAATGTT 6545
 DB 6487 GGATCTGCTGCTTAGGACCTGAGGAAACATGAGGAGTGGGAGCTTCCCTCAATTAACGCT 6546
 QY 6546 ACACGAGGCGCAGTGGCTGCCAAAACCGCGCCAACTTTTAAAGTCCGCTCTGAGGG 6605
 DB 6547 ACACAGGGGCTCTGACTCTCTTCTGCGCGGAACTATAGTTCGCGCTGTTGGAGG 6606
 QY 6606 TGGCGGCTCAGAGTACGGGAGGTGACGAGCAGCGGCTCATACCACTACATTAACAGGAC 6665
 DB 6607 TGTCTGAGAGGAATACGTGGAGATAAGGCGGTGGGGAGCTTCCACTACGATATCGGGTA 6666
 QY 6666 TCACCACTGATTAAGTCCCTGCGCACTACCTCTCCCGAGTCTCTTTTCTCTGGG 6725
 DB 6667 TGACTACTGCAATTTAAATGCGCGTGGCAGATCCCATCGCCGAAATTTTTCACAGAT 6726
 QY 6726 TGACGAGTGCAGATTCATAGGTTTGGCCCCCAGACACGAAAGCGGTTTTTTCGGGATGAGG 6785
 DB 6727 TGAACGGGCTGCGCTTACACAGGTTTGGCGCCCTTGCAGGCCCTTCTGCTGCGGGAGG 6786
 QY 6786 TCTCGTTCTGCGTGGGCTTAATTCATTTGTGCTGGGTCGCCAGCTTCTCTTGGGACCCCTG 6845

Db 2407 TGTGAGTGGTCTTCTCCACCTCCACAAAACATCGTGACGTACAAATTCATGTATGGCC 2466
Qy 2466 TATCACCTGCCCTCACAATAATACATGCTCGATGGAGTGGTAATACCTTATTCCTGC 2525
Db 2467 TATCACCTGCCCTCACAATAATACATGCTCGATGGAGTGGTAATACCTTATTCCTGC 2526
Qy 2526 TCTTAGCGGACGCCAGGTTTGGCCTGCTTATGATGCTCATCTTGTGGGCGAGCGG 2585
Db 2527 TCTTAGCGGACGCCAGGTTTGGCCTGCTTATGATGCTCATCTTGTGGGCGAGCGG 2586
Qy 2586 AAGCAGCATAGAGAGCTGGTCACTTGGACGCTGGAGCGGAGCTAGCTGAATGGCT 2645
Db 2587 AAGCAGCATAGAGAGCTGGTCACTTGGACGCTGGAGCGGAGCTAGCTGAATGGCT 2646
Qy 2646 TCTTATATTTTGTGATCTTTTTCGCTGGCTTGGTACATCAAGGGTGGGCTAGTCCCT 2705
Db 2647 TCTTATATTTTGTGATCTTTTTCGCTGGCTTGGTACATCAAGGGTGGGCTAGTCCCT 2706
Qy 2706 TAGCTACCTATTCCTCAGTGGCTGCTGGTCTTTAGCCTTACTGCTTCTAGCATTGCCTC 2765
Db 2707 TAGCTACCTATTCCTCAGTGGCTGCTGGTCTTTAGCCTTACTGCTTCTAGCATTGCCTC 2766
Qy 2766 AACAGGCTTATGCTTATGACGATCTGTGATGCTGATGAGCGGCTCTGCTGGTAA 2825
Db 2767 AACAGGCTTATGCTTATGACGATCTGTGATGCTGATGAGCGGCTCTGCTGGTAA 2826
Qy 2826 TGAATCACTCTTATCTCACTCACTCCCGGGTATGAGACCTTCTCAGCCGGTTTGTGGT 2885
Db 2827 GGTAAATGGCGCTGACTCTGTGCGCCATATTTACAGCGCTATATCAGCTGGTGCATGGT 2886
Qy 2886 GGTGTGCTATCTTCTGACCTCGGGGAGCTATGCTCAGAGTGGGCAACCACTATGC 2945
Db 2887 GGTCTCAGTATTTCTGACAGAGTAGAGCGCACTGACAGTGGTGGTTCCTCCCTCA 2946
Qy 2946 AGGTGCGCGTGGCGGTGATGATCATATGAGCGCTGCGCATATTTACCCAGGTTGG 3005
Db 2947 ACCTCGCGGGGGCGGATGCGCTCATCTTACTCATGTGTGTAGTACACCGGACCTGG 3006
Qy 3006 TGTTCGACATAACAGTGGCTCTTGGCGGTCTTGGCGTCTTACCTCCTAAAGGTG 3065
Db 3007 TATTTGACATCACCAAACTACTCTGGCCATCTTCGAGCCCTTTGGATCTTCAAGCCA 3066
Qy 3066 CTTTGACCGCGTGGCGTACTCTGCTGAGGCTCACGCTCTACTGAGGATGTGCACCATGG 3125
Db 3067 GTTTGCTTAAGTCCCTACTTCTGCGCGTTCAGGCGCTTCTCCGATCTGCGCTAG 3126
Qy 3126 CAAGGATCTCGCGGGGGGAGTACGCTCAGATGCGGTACTAGCCCTTGGCAGGTGGA 3185
Db 3127 CGCGGAAGATAGCGGAGGTCAATAGTGCATATGCGCAATGCGCATCAAGTTAGGGGCGCTTA 3186
Qy 3186 CTGGCACTTATATGACCACTCACCCCTATGTGGATGGGCTGCTAGTGGCTGC 3245
Db 3187 CTGGCACTTATGATATACCACTCTCACCCCTCTTCGAGACTGGGGCGCAACCGGCTGC 3246
Qy 3246 GGGACCTGGCGGTGGCGTATCATCTTCACTGCTCGATGGAGAAAGTCAATG 3305
Db 3247 GAGATCTGGCGGTGGCTGTGGAAACAGTGTCTTCTCCGATGGAGAACCACTCATCA 3306
Qy 3306 TCTGGGAGCGGAGACAGTCTTGTGGGACATTTTACCGGACTTCCCGTGTCCGCC 3365
Db 3307 CGTGGGGGCGAGATACCGCGCGTGGCTGACATCATCAAGGCTTGGCCGCTCTCGCC 3366
Qy 3366 GACTTGGTGGGAGTCTCTTGGCCAGCTCATGCTATACCTCAAGGGGTGAGTTC 3425
Db 3367 GTAGGGGCCAGGAGATATCTTGGGCCAGCGGATGGTCTCCAGGGGTGGAGT 3426
Qy 3426 TTCTCGCCCATCAGTCTTACCGCCAGCAGACAGTGGCTTTTGGGCAACCATAGTGG 3485
Db 3427 TGTGGCGCCCATCAAGCGGTAGCCCGACAGACAGAGGCTCTTAGGGTGTATATCA 3486
Qy 3486 TGAGCATGACGGGGCGGACAGACAGACAGGCTGGGAAATTCAGGTCTGTCCACAG 3545
Db 3487 CCAGCCTGACTGGCGGGGACAAAACCAAGTGGAGGTTGAGTCCAGATCGTGTCACTG 3546

Qy 3546 TCACTCAGTCTCTCTCGGAACATCCATCTCGGGGTTTTGTGGACTGTCTACCATGGAG 3605
Db 3547 CTACCCAAACCTTCTCTGCAACGTCATCAATGGGTATGCTGGACTGTCTACCAAGGG 3606
Qy 3606 CTGGCAACAGACTCTGGCCGGCTCACGGGGTGGGTTCAGCAGATGCTTCCAGTGTG 3665
Db 3607 CCGGAACAGGACCATCGCATCACCAAGGTTCTGTCTCATCAGATGTATACCAATGTGG 3666
Qy 3666 AGGGGAGCTTAGTGGGTGGCCAGCCCTCGGAGCTAAATCTTTGGAGCGGTGACGT 3725
Db 3667 ACCAAGACTTGTGGGTGGCCCGCTCTCAAGGTTCCCGCTCATTTGACACCCCTGACCT 3726
Qy 3726 GTGAGCGGTGAGCTGTACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3785
Db 3727 GCGGCTCTCTCGGACCTTTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3786
Qy 3786 GCGGGGACAAACGGGGAGGCTTACTCTCCCGAGACCTCTTTTCCACCTTTGAAGGGTCT 3845
Db 3787 GAGGTGATAGCAGGGGTAGCTGCTTTTCCCGCGCCCAATTTCTTACTTTGAAAGGCTCT 3846
Qy 3846 CAGGAGCGCGGTGCTATGCTCCAGGGGCGAGCTGTGCGAGTCTTCCGGGCGCTGTGT 3905
Db 3847 CCGGGGGTCCGCTGTGTGCTCCCGGAGCACGCGGTGGGCTATFTCAAGGGCGCGGTGT 3906
Qy 3906 GCTCTCGGGGCTGGCTAAGTCTCATAGATTTTATCTCCCGTTGAGACACTCCGACATCGTCA 3965
Db 3907 GCACCGGTGGAGTGGCTAAGGGTGGACTTTATCTCTGTGGAGAACCTAGGACACCA 3966
Qy 3966 CCGGCTCCCCCACTTTTATGACACACACCACTGCTGTGTGCTGCTGCTGCTGCTGCTGCT 4025
Db 3967 TGAGATCCCCGGTGTTCACGGACAACCTCTCTCCACAGCAGTGTGCTGCTGCTGCTGCT 4026
Qy 4026 TCGGCTACTTGCATGCTCCCGAGTGGCAGTGGAGAGACACCAAGTTCCTGCTGCTATG 4085
Db 4027 TGGCCCACTGCTGCTCTCCCGGAGCGGTAAAGAGCACCAGGTCCCGGCTGCTGCTGCT 4086
Qy 4086 CTGCTCAGGGGTATAAAGTGTAGTGTCTTATCTCCTCAGTGGCTGCCACCTCGGGTGTG 4145
Db 4087 CAGCCCAAGGCTCAAGGTGTGCTGCTCAACCTCTGCTGCTGCTGCTGCTGCTGCTGCT 4146
Qy 4146 GGGGCTATGCTTAAAGGCAATGCTATCAATCCCAATGAGGCTGAGTCAAGTCAAGTCA 4205
Db 4147 GTGCTTACATGTCCAAAGGCTGCTGCTTATCTTATATCAGGACCGGGGTGAGAACAA 4206
Qy 4206 TGACGACCGGGGCGGCTACGCTACTCCACATATGGCAATTCCTCCCGATGGGCT 4265
Db 4207 TTACCACTGGCAGCCCTACGTAACCTACGCAAGTTCCTTCCCGACGGGCT 4266
Qy 4266 GTGCGGGGCGGCTTACGACATCATATGTGATGAATGCCATGCCGTGGACTCTACCA 4325
Db 4267 GCTCAGGAGTGTCTTATGACATATATTTGTGACAGTGCCTCAGCGGTGCCAT 4326
Qy 4326 CCATCTTTGGCATCGGAACAGTCTTTGATCAAGCAGAGACAGTGGGGTCAAGTCACTG 4385
Db 4327 CCATCTTGGCATCGGCACTGTCTTACCAAGCAGAGACTGCGGGGCGAGACTGGTGT 4386
Qy 4386 TGTGCTTACAGTACGGCCCTGGTCAAGTGAACACCCCGACCCCAACATAGAGAGG 4445
Db 4387 TGTGCTGCACTGCTACCCCTCGGGCTCCGCTCCTGCTGCTGCTGCTGCTGCTGCTGCT 4446
Qy 4446 TGGCCCTTTGGGAGGCGGAGATCCCTTCTATGGAGGGGCTATCCCTGTCTTACA 4505
Db 4447 TGTCTGTCTCCACCGAGAGATCCCTTTTACGGCAAGGCTATCCCTCTCGAGTCA 4506
Qy 4506 TCAAGGGAGGAGACATCTGATCTTTTCTGCTTCAATTAAGAAAGTGTGACAGTCTGCGG 4565
Db 4507 TCAAGGGAGGAGACATCTGATCTTTCTGCTTCAAGAAAGTGTGACAGTCTGCGG 4566
Qy 4566 CCGCCCTTGGGGTATGGCTTGAATCAGTGGGCTACTACAGAGGTTGGAGCTCTCG 4625
Db 4567 CGAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4626

QY 246 CTGCTAGCGAGTAGCGTTGGGTTGGAAAGGCGCTTGTGTACTGTGCTGATAGGGTGTCT 305
 DB 247 CTGCTAGCGAGTAGTGTGGGTGCGAAAGGCGCTTGTGTACTGTGCTGATAGGGTGTCT 306
 QY 306 GCGAGTGCCCGGGAGGTCTCTGAGACCGTGCACATAGAGCAAAATCTTAAACCTCAAA 365
 DB 307 GCGAGTGCCCGGGAGGTCTCTGAGACCGTGCACATAGAGCAAAATCTTAAACCTCAAA 366
 QY 366 GAAAGAACCAAGAAACCAACCGTGCACCAAGAGCGTTAAGTTTCCGGGGCGGGCC 425
 DB 367 GAAAGAACCAAGAAACCAACCGTGCACCAAGAGCGTTAAGTTTCCGGGGCGGGCC 426
 QY 426 AGATCGTTGCGGAGTATATCTTGTGCGCGAGGGGCCCGAGGTTGGGTGTGCGGCGGA 485
 DB 427 AGATCGTTGCGGAGTATATCTTGTGCGCGAGGGGCCCGAGGTTGGGTGTGCGGCGGA 486
 QY 486 CAAGGAAGATTCTGGAGCGTCCAGCAAGTGGAGAGGCGCGAGCCATCCCTTAAAGATC 545
 DB 487 CAAGGAAGATTCTGGAGCGTCCAGCAAGTGGAGAGGCGCGAGCCATCCCTTAAAGATC 546
 QY 546 GCGCTCCACTGCGCAATCTGGGGAACCAAGATACCCCTGCGCCCTATACGGGAATG 605
 DB 547 GCGCTCCACTGCGCAATCTGGGGAACCAAGATACCCCTGCGCCCTATACGGGAATG 606
 QY 606 AGGAGCTCGGCTGGGAGGATGCTCTGCTGCCCGAGGTTCCCGTCTTGGGGCC 665
 DB 607 AGGAGCTCGGCTGGGAGGATGCTCTGCTGCCCGAGGTTCCCGTCTTGGGGCC 666
 QY 666 CCAATGACCCCGGATAGTTCGGGCAACCTGGGTAAAGTTCATCGATACCCCTTAAAGTTCGG 725
 DB 667 CCAATGACCCCGGATAGTTCGGGCAACCTGGGTAAAGTTCATCGATACCCCTTAAAGTTCGG 726
 QY 726 GCTTTGCGGACCTATGCGGTGATATCTGCTGGGCGCGCCGCTCGGGCGGTGCGCA 785
 DB 727 GCTTTGCGGACCTATGCGGTGATATCTGCTGGGCGCGCCGCTCGGGCGGTGCGCA 786
 QY 786 GAGCTCTCGGCGATGCGGTGAGAGTCTCGGAGGAGGGGTAAATTTTGAACAGGGAACT 845
 DB 787 GAGCTCTCGGCGATGCGGTGAGAGTCTCGGAGGAGGGGTAAATTTTGAACAGGGAACT 846
 QY 846 TACCGGTGTCTTTTCTATCTTCTGTTGCTGGCCCTGCTGTCTGATCAACACCCCGG 905
 DB 847 TACCGGTGTCTTTTCTATCTTCTGTTGCTGGCCCTGCTGTCTGATCAACACCCCGG 906
 QY 906 TCTCGCTGCGGAGTCAAGAACATCAGTACCGGCTACATGCTAAACGACTGACCA 965
 DB 907 TCTCGCTGCGGAGTCAAGAACATCAGTACCGGCTACATGCTAAACGACTGACCA 966
 QY 966 ATGACAGCATTACCTGCGAGCTCCAGGCTGTGCTCTCCAGCTCCCGGGTGGGTCCCGT 1025
 DB 967 ATGACAGCATTACCTGCGAGCTCCAGGCTGTGCTCTCCAGCTCCCGGGTGGGTCCCGT 1026
 QY 1026 GCGAGAAAGTGGGAAATGCAATCTCAGTGTGGAATACCGGCTACCGGAATGTGGCCGTGC 1085
 DB 1027 GCGAGAAAGTGGGAAATGCAATCTCAGTGTGGAATACCGGCTACCGGAATGTGGCCGTGC 1086
 QY 1086 AGCGGCGCGGCGCTCAGCAGGCTGCGGAGCATGCAATGCTGATGTCG 1145
 DB 1087 AGCGGCGCGGCGCTCAGCAGGCTGCGGAGCATGCAATGCTGATGTCG 1146
 QY 1146 CCAGCTCTGCTGCTGCTCTACGTTGGGAGCCTCTGCGGTGGGGTGTGCTGCGAGCCC 1205
 DB 1147 CCAGCTCTGCTGCTGCTCTACGTTGGGAGCCTCTGCGGTGGGGTGTGCTGCGAGCCC 1206
 QY 1206 AAATGTTTCAATGCTCGCGGAGCAACATGCTGTTTGTGCAAGCTGCAATGCTCCATCT 1265
 DB 1207 AAATGTTTCAATGCTCGCGGAGCAACATGCTGTTTGTGCAAGCTGCAATGCTCCATCT 1266
 QY 1266 ACCCTGTGATCATCATGCAACCGCATGCGGAGCATGATGCACTGCTGCGCCCA 1325
 DB 1267 ACCCTGTGATCATCATGCAACCGCATGCGGAGCATGATGCACTGCTGCGCCCA 1326
 QY 1326 CCGCTACCATGATCTTGGCGGTACGCGATGCGGTGCTGCCGAGGTCATTTATAGACATCATTA 1385

DB 1327 CCGCTACCATGATCTTGGCGGTACGCGATGCGTGTCCCGAGGTCATTTATAGACATCATTA 1386
 QY 1386 GCGGGCTCATTTGGGCGGTCAATGTTGGCTTGGCTTCTCTATGAGGAGCGGTGG 1445
 DB 1387 GCGGGCTCATTTGGGCGGTCAATGTTGGCTTGGCTTCTCTATGAGGAGCGGTGG 1446
 QY 1446 CGAAAGTCTGTTGCTATCTTCTTGGCCCGCGGGGTGACGCGCACCATATCTGTTG 1505
 DB 1447 CGAAAGTCTGTTGCTATCTTCTTGGCCCGCGGGGTGACGCGCACCATATCTGTTG 1506
 QY 1506 GGGGTTCTCGCGCGAGCAACCGGGCGCTCAAGCTTATTTGACATGAGGCGCCAGCG 1565
 DB 1507 GGGGTTCTCGCGCGAGCAACCGGGCGCTCAAGCTTATTTGACATGAGGCGCCAGCG 1566
 QY 1566 AGAAATCCAGCTCGTTTACCAATGCGAGTGGCACATCAACGCGCCGCTCAACT 1625
 DB 1567 AGAAATCCAGCTCGTTTAAACAATGCGAGTGGCACATCAACGCGCCGCTCAACT 1626
 QY 1626 GCAATGACTCTTGGCACACCGGCTTATCGCGTCTCTGTTTACACCCACAGCTTCAACT 1685
 DB 1627 GCAATGACTCTTGGCACACCGGCTTATCGCGTCTCTGTTTACACCCACAGCTTCAACT 1686
 QY 1686 CGTCAGGATGTCGGAACGATGTCGCTGCGCGAGTATCGAGGCTTTCGGGTGGGAT 1745
 DB 1687 CGTCAGGATGTCGGAACGATGTCGCTGCGCGAGTATCGAGGCTTTCGGGTGGGAT 1746
 QY 1746 GGGGCGCTTGGCAATATAGGATAATGTCAACATCCAGAGATATAGAGCCCTATTTGCT 1805
 DB 1747 GGGGCGCTTGGCAATATAGGATAATGTCAACATCCAGAGATATAGAGCCCTATTTGCT 1806
 QY 1806 GGCACCTACCCACCAAGCGAGTGTGGGTGCTCTCCGCGAGAGTGTGTGTGGGCCCACTGT 1865
 DB 1807 GGCACCTACCCACCAAGCGAGTGTGGGTGCTCTCCGCGAGAGTGTGTGTGGGCCCACTGT 1866
 QY 1866 ACTGTTTACCCCGCGAGTGTGGGTGAGTGGGAGAGTGTGGAGGCTTGGAGCGCCACTTT 1925
 DB 1867 ACTGTTTACCCCGCGAGTGTGGGTGAGTGGGAGAGTGTGGAGGCTTGGAGCGCCACTTT 1926
 QY 1926 ACACGTGGGGAGATGAGACAGATGCTTCTCTATTGAACAGACTCGACACACCGCTGG 1985
 DB 1927 ACACGTGGGGAGATGAGACAGATGCTTCTCTATTGAACAGACTCGACACACCGCTGG 1986
 QY 1986 GGTTCATGTTTGGCTGCAAGTGGATGAACTCTTCTGGCTACACCAAGACTTGGGGCGCAC 2045
 DB 1987 GGTTCATGTTTGGCTGCAAGTGGATGAACTCTTCTGGCTACACCAAGACTTGGGGCGCAC 2046
 QY 2046 CACCTGCGGTACTAGAGTGAATCAACGCGAGCAAGGACTGTTGTGCGCCCAAGACT 2105
 DB 2047 CACCTGCGGTACTAGAGTGAATCAACGCGAGCAAGGACTGTTGTGCGCCCAAGACT 2106
 QY 2106 GTTTTAGGAGGACTCTGATACCACTTACCTCAAAATGCGGCTCTGGGCGCTGCTCAAGC 2165
 DB 2107 GTTTTAGGAGGACTCTGATACCACTTACCTCAAAATGCGGCTCTGGGCGCTGCTCAAGC 2166
 QY 2166 CAAGGTGCTGATGCTGCTACCGCTTGGGCTTACCGCTGCGACAGTAACTATA 2225
 DB 2167 CAAGGTGCTGATGCTGCTTACCGCTTGGGCTTACCGCTGCGACAGTAACTATA 2226
 QY 2226 CCAATTTCAAAATAGGATGATGTTGGAGGGGTGAGGAGGCTCAGGCTGCGATGCA 2285
 DB 2227 CCAATTTCAAAATAGGATGATGTTGGAGGGGTGAGGAGGCTCAGGCTCAGGCTGCA 2286
 QY 2286 ATTTTCACTGTTGGGATCGTTGCAATTTGGAGGAGAGAGAGTCAACTGTCTCTCT 2345
 DB 2287 ATTTTCACTGTTGGGATCGTTGCAATTTGGAGGAGAGAGAGTCAACTGTCTCTCT 2346
 QY 2346 TGTGCTACTCCACCGGATGGGCAATTTTACCTTGTCTTACTCGGACTGCGCGCT 2405
 DB 2347 TGTGCTACTCCACCGGATGGGCAATTTTACCTTGTCTTACTTGTGAGCTGCGCGCT 2406
 QY 2406 TGTGCTACTGCTTCTCACCTCCACAAACATCGTGGAGCGTCAACTCATGTATGCGC 2465

8641 TGACCAAGTACTCGGCGCCCGGCGGACCCCAACAGAAATACGACTTGGAGCTTA 8700
8754 TACATCTTCTCTCAATATGCTCTGTGGCGTGGGCCCAACAGCGCCGCGAGTACT 8813
8701 TAAATCATGCTCTCTCAAGTGTAGTGTGCGCCACAGCGCGCTGGAAGAGGCTTACT 8760
8814 ACCTGACGAGAGACCTTACCACCTCAATCGCCGCGCTGCTGGGAAACAGTTAGACACT 8873
8761 ACCTTACCGTGAACCTTACACCCCGCTCGGAGAGCGCGGTGGGAGACAGCAGACACA 8820
8874 CCCTGTCAATTCATGCTGGGAAACATCATCAGTACGCGCCCGACCATATGGCTCGCA 8933
8821 CTCAGTCAATTCCTGCTAGGCAACATATATGTTGGCCCGCCACACTGTGGCGGAGGA 8880
8934 TGGTCTGTATGACACACTTCTTCTCAATCTCATGCTCAAGACAGCTGGACCAAGCC 8993
8881 TGATACTGTATGACCCCAATTTCTTTAGCGTCTCTATAGCCAGGATCAGCTTGAACAGGCTC 8940
8994 TCAACTTTGAGATGTACGAGCGGTGTACTTCCGTGAGTCCCTTGGACCTCCAGCTTATAA 9053
8941 TTAACGTGTAGATCTACGAGCGCTGTACTTCCATAGAACCACTGGATCTACCTCCAAATCA 9000
9054 TTGAAAGGTTTACATGGCTTACGCTTTTCTCTGCACACATACACTTCCCGACGAACTGA 9113
9001 TTCAAAGACTCCATGGCTCAGCGCTCAGCGCATTTTCACTCCACAGTTACTTCCAGGTGAATCA 9060
9114 CAGGCTGGCTTACGCGCTCAGAAACTTTGGGCGGCCACCCCTCAGAGCGTGAAGAGCC 9173
9061 ATAGGCTGGCGCATGCTTCCAGAAACTTTGGGCTCCCGCTTGGCAGCTTGGAGACAC 9120
9174 GCGCAGCTGAGTCAAGGCTGCTCTCATCTCCCGTGGGGGAGAGCGCGTGTGGCGTC 9233
9121 GGGCGCGGCGCTCCGCGCTAGCGTCTGTCCAGAGAGGCGGCTGTATATGTGGCA 9180
9234 GATATCTTCAATTTGGGCGGTGAAGACCAAGCTCAAACTCACTCAATTCGCGAAGCGC 9293
9181 AGTACCTCTTCAACTGGGCGAGTGAAGAACAAAGCTCAAACTCACTCAATTAGCGCGCTG 9240
9294 GCCTCTGATTTATCAGCTGTTTCAACGCTGGCGCGCGCGGGGCGGAGCATTTATCACA 9353
9241 GCGGCTGGAGTGTGCTCGGTTGTTTCAAGCTGCTTACAGCGGGGAGACATTTATCACA 9300
9354 GCGTGTGCGGTGCGGCGCGCTTATGCTCTTTCGCGCTACTCTTCTTGTAGGGG 9413
9301 GCGTGTCTATGCGCGCGCGCTGCTTCTGTTTTCGCTTCTTCTGCTGCTGCGAGGG 9360
9414 TAGGCGCTTTTCTTACTCCCGCTCGGTAGAGCGGACACATTAGCTACACTTCCATAGCTA 9473
9361 TAGGCATCTACTCTCTCTCCCAACCGATGAAGTTGGGGTAAACACTCCGCGCTCTTAAAGC 9420
9474 ACTGTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9533
9421 CATTTCTGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9480
9534 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 9593
9481 TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 9497
9594 CTTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 9653
9498 ---TTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 9553
9654 GGTGCGTGAAGCGCATGCTGAGAGAGTGCCTAATCTGCTCTCTCTGAGATCATGT 9711
9554 GGTGCGTGAAGCGCATGCTGAGAGAGTGCCTAATCTGCTCTCTCTGAGATCATGT 9611

RESULT 8
AAC86647
ID AAC86647 standard; DNA; 9611 BP.
XX
AC AAC86647;

XX 02-APR-2001 (first entry)
XX Nucleotide sequence of chimeric Hepatitis C virus clone J6S.
XX HCV; HCV strain HC-J6CH; HCV genotype 2a; antiviral; vaccine; ss.
XX Synthetic.
XX Hepatitis C virus.
XX Key Location/Qualifiers
XX CDS 342..9389
XX /*tag= a
XX WO200075338-A2.
XX 14-DEC-2000.
XX 02-JUN-2000; 2000WO-US15446.
XX 04-JUN-1999; 99US-0137693.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX Yanagi M, Bukh J, Emerson SU, Purcell RH;
XX WPI: 2001-061728/07.
XX P-PSDB; AAB30732.
XX Nucleic acid molecule encoding human hepatitis C virus of genotype 2a
XX for developing vaccines, for diagnosis of hepatitis C virus and in
XX screening assays for identification of antiviral agents -
XX Disclosure; Page 131-134; 167pp; English.
XX AAC86645-48 represent chimeric cDNA clones of infectious Hepatitis C
XX virus (HCV). In each clone the C, E1 and E2 genes are derived from HCV
XX strain HC-J6CH genotype 2a. The p7 protein was derived from HCV strain
XX HC-J6CH or HCV strain pcv-H77C, and the NS genes were derived from HCV
XX strain pcv-H77C. Such HCV sequences are capable of expressing the virus
XX when transfected into cells. The HCV protein is useful for assaying
XX candidate antiviral agents for activity against HCV. Antibodies specific
XX for HCV polypeptide are useful in prevention and treatment of diseases
XX caused by HCV in animals, in particular humans. The HCV polypeptides
XX serve as immunogens in the development of vaccines for preventing HCV
XX in mammals or as antigens in diagnostic assays for detecting the
XX presence of HCV in biological samples. The HCV polynucleotide is also
XX useful for identifying cell lines capable of supporting the replication
XX of HCV in vitro and to produce attenuated viral strains via passage in
XX vitro or in vivo.
XX Sequence 9611 BP; 1935 A; 2896 C; 2682 G; 2098 T; 0 other;
XX
XX Query Match 59.7%; Score 5800.8; DB 22; Length 9611;
XX Best Local Similarity 75.4%; Pred. No. 0;
XX Matches 7326; Conservative 0; Mismatches 2267; Indels 125; Gaps 4;
XX
XX 6 CCCCTAATAGGGCGAGACTCGGCATCAATCACTCCCTGTGAGGAATCTGCTTCA 65
XX 7 CCCCTGATGGGGCGGACACTCCACCATGAATCACTCCCTGTGAGGAATCTGCTTCA 66
XX 66 CGCAGAAAGCGTCTAGCCATGCGTCTAGTATGAGTGTCTGACAGCTCCAGGCCCGCC 125
XX 67 CGCAGAAAGCGTCTAGCCATGCGTCTAGTATGAGTGTCTGACAGCTCCAGGCCCGCC 126
XX 126 TCCCGGGAGAGCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTCGCGGAAGAC 185
XX 127 TCCCGGGAGAGCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTCGCGGAAGAC 186
XX 186 TGGGTCTCTTCTTGGATAAACCCACTCTATGCCCGGCCATTTGGCGCTGCCCGCAAGA 245
XX 187 CGGGTCTCTTCTTGGATAAACCCCGCTCAATGCTGAGATTTGGGCGTGGCGCCGCAAGA 246

QY	6546	ACACGAGGCGCAGTGGTCCGAAACCCGCGCAAACTTTAAAGTTCGCATCTGGAGG	6605	7614	ACTCGGGTCTCTGTTACTTCTCCGAGGAGGACGACTCCGTCGTGTGCTCTCAATGT	7673
Db	6547	ACACAGGCGCCCTGTATCTCCCTCTCTGCGCGCACTATAAGTTTCGCGCTGTGAGG	6606	7561	GGTCATGTGCGAGCGTCACTAGTGGGCGCACACGGAAGATCTGCTGCTCAATCT	7620
QY	6606	TGCGGCGCTCAGAGTACGCGAGGTGACGAGCAGCGGTTCATCACTACATAACAGGAC	6665	7674	CATATCTCTGGACCGGGCTCTAAATCTCTGTAGTCCGCAAGGAGGAGTTACCGA	7733
Db	6607	TGTCGAGAGGAATACGTGAGATAAGCGGTGGGAGATTCACACTACGTATCGGTA	6666	7621	CTTATTTCTGGACAGCGCACCTCGTACCCCGTGGCTCGGGAAGAACAAAACCTGCCCA	7680
QY	6666	TCACCACTGATTAATGAAAGTCCCTGCGCAACTACCTCTCCGAGTCTTTCTCTGG	6725	7734	TTAAACCCCTTGAGCAACTCCCTGTTGCGATATACAAACAGGTGTACTGTACCAACAA	7793
Db	6667	TGACTACTGACAACTTAAATGCGGTGCGAGATCCCATCGCCGAAATTTTTCACAGAT	6726	7681	TCACCGCACTGAGCAACTCTGTTGCTAGCGCATCACAACTGTTGTTATTCACCACTTCA	7740
QY	6726	TGACCGAGTGCAGATTCATAGTGTTCGCCCCACACCGAAGCGGTTTTTTCGCGATGAGG	6785	7794	AGAGCGCTTCACTAAGGGCTTAAAGGTAACTTTTGTATAGGATGCAAGTCTCGACTCT	7853
Db	6727	TGACCGGGTTCGCCCTACACAGTTTGGCGCCCTTTCAGAGCCCTTCTGCGGGAGG	6786	7741	GCAGTCTTTCGCAAGGCAAGAAAGTCACTTTGACAGACTGCAAGTTCTGGACAGC	7800
QY	6786	TCTGTTCTCGTGTGGCTTAAATCACTTGTGTCGCGTCCAGCTTCTTTCGAGCCCTG	6845	7854	ACTAGCACTCACTTAAAGGACATTAAGCTAGCGCTCCCAAGGTTCACGCAAGCTCC	7913
Db	6787	TATCAITTCAGAGTAGGACTCCAGAGTACCGGTGGGTGCGCAATTACCTTTCGAGCCG	6846	7801	ATTACCGAGACGTCTCAAGGAGGTCAAGCAGCGCGTCAAAAGTGAAGGCTTAATTGC	7860
QY	6846	AACCGACACAGAGTATTCATGTCATGCTTAACAGATCCATCTCATATCACGCGGAGA	6905	7914	TCACCATGAGGAGGCTTGCAGTTAAACCCCACTTCTGCAAGATCTAAATATGGGT	7973
Db	6847	AACCGGACGTAGCGGTTCACGTCTCATCTGATCCCTCCCATATAACAGCAGAGG	6906	7861	TATCCGTAGAGGAGCTTGCAGCTGACGCCGCCCACTTCAGCCAAATCCAGTTTGGCT	7920
QY	6906	CTGCAGCGCGGTTAGCGCGGGGTCAACCCCATCCGAGGCAAGCTCTCTCGCGAGCC	6965	7974	TTGGGCTTAAGGAGTCCGAGCTTGTTCGGGAGGCGCTTAAACCACTCAAGTCCGTGT	8033
Db	6907	CGCGCGGAGAGGTTGCGAGAGGTTACCCCTCTCTATGCGCAGCTCTCTCGGTAGCC	6966	7921	ATGGGCAAAAGACGCTTCGTCATGCCAGAAAGCGCTAGCCACATCAACTCCGTGT	7980
QY	6966	AGCTATCGGCACCATCGCTCGAGCCACCTGCAACCAACCGCAAGCTTATGATGTGG	7025	8034	GGAGGACCTCTTGGAGGACTCAGAAACACCAATTTCCCAACCACTTATGGCCAAAGG	8093
Db	6967	AGCTGTCGCTCCATCTCTCAAGSCAATTTGCAACCGCAACCATGACTCCCTGACGCG	7026	7981	GGAAAGACCTTCTGGAAGACAGTGTAAACCAATAGACACTACCATCATGGCCAGACG	8040
QY	7026	ACATGTGATGCTAACT-----GTTTCATGGGGCGATGTGACTCGATAG	7073	8094	AGTGTCTTCTGCTGGACCCCAACAGGGGGGCAAGAAAGCAGCTCCCTTATCGTTTACC	8153
Db	7027	AGCTCATAGAGGTAACCTCTCTGAGGAGGAGATGGCGGCAACATCACAGGGTTG	7086	8041	AGGTTTCTGCTTCACTGAGAGGGGGTTCGTAAGCCAGCTCTCTCACTCGTCTCC	8100
QY	7074	AGTCTGGTCCAAAGTGGTCTCTGACCTCTCTGACCAACCTTGTGCAAGAAAGGAGCG	7133	8154	CTGACCTCGCGCTCAGGCTTTCGAGAGATGGCCCTTTATGACATATACAAAACTTC	8213
Db	7087	AGTCAGAGAACAAAGTGTGATCTGACCTCTCTGATCGCTTGTGGCAGAGGATG	7146	8101	CGACCTGGCGTGGCGTGTGCGAGAGATGGCCCTGTACGACGTGGTTAGCAAGTCC	8160
QY	7134	ACCTTGAGCTTCGATACCAATCAGATPACATGTCTCCCAAGAGAGGTTCCCAACAGCTT	7193	8214	CTCAGCGGTGATGGGGCTTCTTATGGATTCAGTATTTCCCGCTCAGCGGGTAGAGT	8273
Db	7147	AGCGGAGGTCTCGTACCTGCGAATTTCTGCGAAGTCTCGAGATTCGCGCGGCC	7206	8161	CCCTGGCGGTGATGGGAAGCTCTACGATTCCTCAATCTACCCAGGACGCGGTGAT	8220
QY	7194	TACCGGCTGGCAAGCGCTGATTAACCCACCGCTTGTGATCGTGAAGAGCCAG	7253	8274	TTCTCTTAAAGCATGGGGGCAAGAACGACCTATGGGTTTTTCTGATGATACCCGAT	8333
Db	7207	TGCGCGCTGGCGCGCGGACTACAAACCCCGCTAGTAGAGACGTGGAAGAGCTG	7266	8221	TCCTCTGCAAGCGTGGAGTCCAAGAGACCCGATGGGTTCTCGTATGATACCCGCT	8280
QY	7254	ATTACCAACCGGCACCTGTTGCGGCTGTCTCTCCCTCTCTAGGAAAAACCCGACGC	7313	8334	GTCTTGACTCAACCGCTCACTGAGAGAGACATCAGGACTGAGGAGTCCATATCGGGCCT	8393
Db	7267	ACTACGAACACCTGTGTCATGGCTGCGGCTACCACTCCACGCTCCCTCTCTGTGC	7326	8281	GTCTTGACTCAACGTCATCTGAGAGGACATCCGTACGAGGAGGCAATTTTACCAATGTT	8340
QY	7314	CTCCCCAAGAGGCGCGGACAGTGGGCTTAAGTGAAGACTCCATAGGAGATGCCCTTC	7373	8394	GCTCTTGGCGAGGAGGCGCCACACTGCCATACACTCGCTAACTGAGAGACTTTTACGTGG	8453
Db	7327	CTCGCTCTGGAAGAGCGTACGCTGCTCTACCGAATCAACCTATCTACTGCTTGG	7386	8341	GTGACCTGGAACCCCAAGCGCGGTGGCCATCAAGTCCCTCACTGAGAGGCTTTATGTTG	8400
QY	7374	AACAGCTGGCATTAAGTCTTTGGCCAGCCCCCCCCCAAGCGGATTCAGGCTTTCCA	7433	8454	GAGGCTTATGTTCAACAGCAAGGCGCAACCTTCGCGGTACAGGCGTTCGCGCGCCAGCG	8513
Db	7387	CCGAGCTTGCCACCAAAAGTTTGGCAGCTCTCAACTTCGCGCATTTACGGGCGCAATA	7446	8401	GGGGCTCTTTACCAATTCAGGGGGGAAATCTGCGGCTACCCAGGTCCGCGCGAGCG	8460
QY	7434	CGGGGCGGGCTGCGGATTCGCGGAGTCAGAGCGCTCTCTGATGATGTGGCCCTTCGG	7493	8514	GGGTGCTCACTAGCATGGGGAACCACTACATCATGCTACGTAAGTGAAGCTTTAGCGGCTT	8573
Db	7447	CGACAACTCTCTG-----AGCCCGCCCTTCTGCGTCCGCGCGCCGAGCTCG	7494	8461	GCCTACTGACAACTAGCTGTGGTAAACCCCTCACTTGTCTACATCAAGGCGCGGAGCCT	8520
QY	7494	AGACAGGTTCCATCTCTTCCATGCCCCCTTCGAGGGGAGGCTTGGAGATCCAGACCTGG	7553	8574	GTAAAGCTGAGGGATTAATCGCGCCCACTATGCTGTTATGCGGAGTACTTGTGTTCA	8633
Db	7495	ACGTTGAGTCTATTCTTCCATGCCCCCTTCGAGGGGAGGCTTGGGATCCGATCT--	7552	8521	GTGAGCGCGAGGGCTCCAGGACTGACCACTGCTGCTGTGTGGCGAGCACTTAGTCTGTTA	8580
QY	7554	AGCCTGAGAGGTAGAGCCCCCAACCCCCCCCCCAGGGGGGGTGGAGCTCCCGGCTCGG	7613	8634	TCTCAGAAACCGCAGGAGGACCGAGGAGGACCGGCGGAGCTTCAGAGGCTTCAACGAGGCTA	8693
Db	7553	-----CAGCGAGC	7560	8581	TCTGTGAAGTGGCGGGGTTCAGAGAGGACCGGCGGAGCTTCAGAGGCTTCAACGAGGCTA	8640
				8694	TGACCAGGTTATCTGCCCCCTCTCTGGTGAACCCCCCCCCCAGACCGGAGTATGATCTGGAGCTGA	8753

5407	GGATCGTCTTGTCGGGGAAGCCGGCAATTA	TACCTGACAGGAGGTTCTCTACCAAGAGT	5466
5466	TTGATGAGATGGAGGAATGTGCCTCTAGGCGGGTCTCATTTGAAGAGGGGAGCGGATAG	5525	
5467	TCGATGAGATGGAAGAGTGTCTCAGCACTTACCGTACATCGAGCAAGGATGATGCTCG	5526	
5526	CCGAGATGCTGAAGTCCAAGATCCAAGGCTTATTGACGAAAGCTTCCAAA	5585	
5527	CTGAGCAGTTCAAAGCAGAAGGCCCTCGGCGCTCTCTGCAGACGCGTCCCGCCATGCAGAGG	5586	
5586	ACATACAAACCACATCTGTGACAGCTTTCATGGCCCAAGGTAGAACNAATCTGGGCCCCAACAACA	5645	
5587	TTATCACCCCTGCTGTCCAGACCAACTGGCAGAAACTCGAGGTCTTTTGGGCGCAAGCACACA	5646	
5646	TGTGGAACTTCATTAGCCGGCATCCAATACCTCGCAGGACTATCAACACTGCCAGGGAACC	5705	
5647	TGTGGAACTTCATCAGTGGGATACATATCTTGGCGGGCTGTCAACGCTGCTGTGTAACC	5706	
5706	CTGCAGTAGCTTCCATGATGCGGTTCAAGTCCGCGCTCACAGTCCGCTGTCAACAAGCA	5766	
5707	CGGCCATTGCTTCATTGATGCTTTTACAGCTGCGCTCACCAAGCCCACTAACCACTGGGCC	5766	
5766	CCACTATCTTCTCNAACATTTTGGGGGCTGGCTAGCATCCCAATTCGACCAACCCCGGG	5825	
5767	AAACCCCTCTCTTCAACATAATTTGGGGGGTGGGTGGCTGCCAGCTCGCCGCCCGCGGTG	5826	
5826	GGGCCACTGGCTTCGTTCTCAGTGCCTPAGTGGGAGCTGCCGTAGGCAAGTATAGCTTTAG	5885	
5827	CGCTACTTGCCTTTTGTGGGTGCTGGCTTGGCTTGGCGCGCCATCGCAGCGTTGGA	5886	
5886	GTAAGGTGCTAGTGGACATCCTCGCAGGGTATGTTGCGGGCAATTTTCGGGGCTCTCGTCTG	5945	
5887	GGAAGTCTCTGTGGACATTTCTTCAGGGTATGGCGGGCGTGGCGGGAGCTCTTGTAG	5946	
5946	CATTCAAGATCATGTTCTGGGAGAAAGCCCTCCATGGAGGATGTGTTCAACTTGTGCTGCTG	6005	
5947	CATTCAAGATCATGAGCGGTGAGGTCCCTCCACGGAGGACCTGGTCAATCTGTCTGCCCG	6006	
6006	GAATTCGTCTCTCGGGTGCTTGTGTAGTGGGAGTCACTGTGCGGCGCAATCTGCGGCCGAC	6065	
6007	CCATCTCTCGCTTGGAGCCCTTGTAGTTCGGTGTGTTCTGCGAGCAATCTGCGCCGCGC	6066	
6066	ACGTGGGACCGGGGGAAGCGCGCTCCAATGGATGAATAGACTCATTTGCCCTTTGCTTCCA	6125	
6067	ACGTTGGCCCGCGGAGGGGCGAGTGCATATGATGAACCGGCTAATAGCTTTCGCTCC	6126	
6126	GAGGAATCACTGCGCCCCACCACTACGTGACGAGTTCGGATCGCTCGCAGCGTGTGA	6185	
6127	GGGGGAACCATGTTTTCCCCACGCACTACGTGCGGAGAGGAGTGACGCCCGCGCTCA	6186	
6186	CCCAACTACTTGGCTCCCTTACCATACACGCTGCTCAGAAGACTCCACAACCTGGATT	6245	
6187	CTGCCATCTCAGCAGCCTCACTGTAAACCCAGCTCTTGAGGCGACTGCATCAGTGGAT	6246	
6246	CTGAGGACTGCCCATCCCATGCGCGGCTCGTGGCTCCGCGATGTGTGGACTTGGGTTT	6305	
6247	GCTCGGAGTGTACCACTCCATGCTCCGGTCTCTGGCTAAGGGAATCTGGGACTG	6306	
6306	GCACCATCTTAACAGACTTTAAATTTGGCTGACCTCCAAATTAATTC	6365	
6307	CGGAGGTGCTGAGCGACTTTAAGACCTTGGCTGAAAGCCCAAGCTCATGCGACA	6366	
6366	GCCTCCCTTTGTCTCTGTCAAAGGGGTACAGGGGCTGTGGCCGGCACTGGCATCA	6425	
6367	GGATTCCTTTGTCTCTGCCAGCGCGGTATAGGGGGGTCTGGCGAGGAGACGCGCAT	6426	
6426	TGACCACACGGTGTCTTGGCGGCCAATATCTCTGGCAATGTCCGCTTGGCTCCATGA	6485	
6427	TGCACATCTCGTCCCATGTGTGAGCTGAGATCATCTGGNCATGTCAAAA	6486	
6486	GAATCACGGGGCTTAAGACTTCATGATGAATATCTGGCAGGGGACCTTTCTTCAAT	6545	
6487	GGATCGTCTGCTTAGGACCTTCAGAGAACATGTGGAGTGGGACGTTTCCCAATTAACGCT	6546	

2166 CRAAGTGGCTGATGACTACCCCTACAGGCTCTGGCAATACCCCTGCACAGTTAACTATA 2225
 2167 CRAAGTGGCTGATGACTACCCCTACAGGCTCTGGCAATACCCCTGCACAGTTAACTATA 2226
 2226 CCATCTTCAAAATAAGATGATATGTGGAGGGGTGAGCAACAGCTCACGGCTGCATGCA 2285
 2227 CCATCTTCAAAATAAGATGATATGTGGAGGGGTGAGCAACAGCTCACGGCTGCATGCA 2286
 2286 ATTTCACTCTGGGGATCGTTGCACTTGGAGACAGACAGCAAGTCAACTCTCTCCTT 2345
 2287 ATTTCACTCTGGGGATCGTTGCACTTGGAGACAGACAGCAAGTCAACTCTCTCCTT 2346
 2346 TGTGTCACTCCACACAGGAATGGCCATTTTACTTGTCTTACTCGGACCTCCCGCCCT 2405
 2347 TGTGTCACTCCACACAGGAATGGCCATTTTACTTGTCTTACTCGGACCTCCCGCCCT 2406
 2406 TGTGCACTGCTTCTTCCACTCCACAAACATCGTGGAGCTGCAATCATGTATGGCC 2465
 2407 TGTGCACTGCTTCTTCCACTCCACAAACATCGTGGAGCTGCAATCATGTATGGCC 2466
 2466 TATCACTGCTCCACAAATACATCGTCCGATGGGAGTGGTAACTCTTATCTCTG 2525
 2467 TATCACTGCTCCACAAATACATCGTCCGATGGGAGTGGTAACTCTTATCTCTG 2526
 2526 TCTTAGCGGACGCGGTTTGGCCCTGCTTATGGATGCTCATCTTGTGGGCGAGCCG 2585
 2527 TCTTAGCGGACGCGGTTTGGCCCTGCTTATGGATGCTCATCTTGTGGGCGAGCCG 2586
 2586 AAGCAGCACTAGAGAGCTGGTCACTTTCGACCGTGGAGCGAGCTAGCTGCAATGGCT 2645
 2587 AAGCAGCACTAGAGAGCTGGTCACTTTCGACCGTGGAGCGAGCTAGCTGCAATGGCT 2646
 2646 TCCTATATTTGTCATCTTTTGGTGGCTGCTTGGTACATCAAGGGTGGGATGCTCCCT 2705
 2647 TCCTATATTTGTCATCTTTTGGTGGCTGCTTGGTACATCAAGGGTGGGATGCTCCCT 2706
 2706 TAGCTACTATTTCCCTCACTGGCTGCTGCTTTAGCTTACTGCTTACTAGCAATGGCC 2765
 2707 TAGCTACTATTTCCCTCACTGGCTGCTGCTTTAGCTTACTGCTTACTAGCAATGGCC 2766
 2766 AACAGGCTTATGCTTATGAGCACTCTGTGATGCGCCAGATAGAGAGCGCTCTCTGTGTA 2825
 2767 AACAGGCTTATGAGCACTCTGTGATGCGCCAGTGGTGGCGGCTGCTTCTGTGTCG 2826
 2826 TGATCACTCTTTACTCTCACCCCGGGTATAGACCTTCTCAGCGGCTTTTGTGCT 2885
 2827 GGTAAATGGCGCTGACTCTGTGCGCATATTACAGCGCTATATCAGCTGGTGCATGCTGT 2886
 2886 GGTGTGCTATCTTCACTGCGGAGTATGGTCCAGGAGTGGGCAACCACTATGCT 2945
 2887 GGCTTCAGTATTTCTGACAGAGTAGAGGCAACTGCACTGTGGGTTCCCGCCCTCA 2946
 2946 AGTGGCGGTGGCGGTGATGGCACTATATGGGCGGTGCGCATATTCTACCAAGGTGG 3005
 2947 ACGTCCGGGGGCGCGATGCGCTCATCTTACTCATGTGTGTAGTACACCCGACCTGG 3006
 3006 TGTGTGACATAACCAAGTGGCTCTGGGCTGCTTGGGCTGCTTACTCTTAAAGGTG 3065
 3007 TATTTGACATACCAAACTACTCTGGCCATCTTTCGAGACCCCTTTGGATTTCTCAAGCCA 3066
 3066 CTTTGAAGCGGTGCGCTTCTGCTCAGGCTCAGCTTACTGAGGATGTGCACCACTGG 3125
 3067 GTTTGCTTAAAGTCCCTACTTCTGTCGCTTCAAGCCCTTCTCCGATCTGCGCGCTAG 3126
 3126 CAAGGCATCTGGCGGGGCGAGGTACGTTCAGATGGCGCTTACTAGCCCTTGGCAGGTGGA 3185
 3127 CGCGGAAGATAGCGGAGGTCAATTACGTGCAATGGCCATCATCAAGTTAGGGGCGCTTA 3186
 3186 CTGGCACTTATCATCTATGACCACTCACCCCTATGTCGGATGGGCTGCTAGTGGCCCTGC 3245
 3187 CTGGCACTTATGTTATTAACCACTCTCCCTCTTCGAGAGCTGGGCGCAACAGCGGCTGC 3246

3246 GGGACCTGGCGCTGCGCTTGGAGCTTATCATCTTCACTCCGATGGAGAAAGTCAATG 3305
 3247 GAGATCTGGCGCTGGCTGGAAACAGTCTTCTTCCCAATGGAGAACAGTCAATCA 3306
 3306 TCTGGGAGCGGAGACAGCTGCTTGGGGAATTTTACACGCACTTCCGCTGTCGCGCC 3365
 3307 CGTGGGGGCAATACCGCGCTGCGGTGACATCATCAACGGCTTCCCGCTCTGCGCC 3366
 3366 GACTTGGTGGAGGCTCTCTTGGCCAGCTGATGGCTTATCTTCAAGGGGTGGAGTC 3425
 3367 GTAGGGCCAGAGATATCTGCTTGGCCAGCCGAGATGCTTCCAGAGGGGTGGAGT 3426
 3426 TTTGCGCCCTCATCTGCTTACGCCAGCAGACAGTGGCTTTTGGGACCATAGTGG 3485
 3427 TGTGGCGCCATCATCGGCTACCGCCAGCAGAGAGCGCTCTTAGGCTGTATATCA 3486
 3486 TGAGCATAGCGGGCGGAGACAGACAGCTGGGGAATTTAGGTCCTGTCACAG 3545
 3487 CCAGCTGATGCGCGGGAACAAACAAAGTGGAGGTGAGGTCCAGATCGTGTCAACTG 3546
 3546 TCACTCAGTCTCTTCTCGGAACATCATCTCGGGGTTTTGTGGACTGTCTACCATGGAG 3605
 3547 CTACCCAAACCTTCTGGCAACGTCATCAATGGGTATGCTGGAATGTCTACCAAGGG 3606
 3606 CTGGCAACAGACTCTGGCGGCTCACGGGTCCGGTCCGAGATGTACTCAGTGTCTG 3665
 3607 CCGAACAGGACCATCGCATCACCAAGGCTCTGCTCATCCAGATGTATACCAATGTGG 3666
 3666 AGGGGACTTACTAGGTTGGCGCCAGCCCTCGGACTAAATCTTTGGAGCGGTGACCT 3725
 3667 ACCAAGCTTGTGGGCTGGCGGCTCTCAAGTTCCCGCTCATTTGACACCTGTACCT 3726
 3726 GTGGAGCGTCACTGTACTGCTCAACGCGAAACGCTGATGTATCTCCGGCTCGAAGAC 3785
 3727 GCGCTCTCGGACCTTACTGCTCAACGAGCAACGCGCTGCTTCCCGTGGCGCGC 3786
 3786 GCGGGACAAAACGGGAGCGCTACTCTCCCGAGACTCTTTCACCTTTGAAGGGTCT 3845
 3787 GAGTGTAGAGGGGTAGCTGCTTTCGCGCCCGGCCATTTCTACTTTGAAGGGTCT 3846
 3846 CAGGAGCGGTGCTTATGCGCCAGGGCCAGCTCTCGGAGTCTTCCGGGAGCTGTGT 3905
 3847 CCGGGGTCTCGCTTGTGCGCGGGACACGCGCTGGGCTTATTCAGGGCGCGGTGT 3906
 3906 GCTCTCGGGCGTGGCTTAACTTCAATCTTCCCGTGGAGCACTCGACATCGTCA 3965
 3907 GCACCGTGGAGTGGCTTAAAGGCTGACTTTATCTCTGTGGAGAACTAGGACACCA 3966
 3966 CCGGCTCCCGCTTGTAGCAACAGCAGCACTCTGCTGTGGCGCCAGCTATCAGG 4025
 3967 TGAGATCCCGGTGTTCAGGCAACTCTCTCCAGCAGAGTGGCGCCAGAGCTTCCAGG 4026
 4026 TCGGCTACTTGTGATGCGCCAGCTGGCAGTGGGAAGAGCAACAAAGTTCCTGTCGCAATG 4085
 4027 TGCGCCACTGCTATGCTCCCGCGGAGCAACGCGCTGGGCTTATTCAGGGCGCGGTGT 4086
 4086 CTGCTCAGGGGTATAAGTGTAGTCTTAACTTCCCTCAGTGGTGGCCACCTCGGGTTG 4145
 4087 CAGCCAGGGCTACAGGTGTGTGCTCAACCCCTCTGTTGCTGCAACGCTGGGCTTG 4146
 4146 GCGGCTACTTGTCTAAGGACATGGCATCAATCCCAACATTAGGACTGGAGTCAAGCTG 4205
 4147 GTGCTTACATGTCCAAGGCCATGGGTTGATCTTAATATCAGGACCGGGGTGAGAACAA 4206
 4206 TGACGACCGGGCGCCCATCAGTACTCCACATATGCGAAATTCCTCGCGATGGGGCT 4265
 4207 TTACCACTGGACCCCATCAGCTACTCCACCTTACCGCAAGTTCCTTTCGCGAGCGGGT 4266
 4266 GTCGGCGCGCTACGACATCATATATGTGATGAATGCCATGCGCTGGAATCTTACCA 4325
 4267 GCTCAGAGGTGCTTATGACATATATTTGTGACAGTGGCACTCCACGATGCCACAT 4326
 4326 CCATCTTGGCATCGGAACAGTCTTGTATCAAGCAGACAGCTGGGGTCAAGTAACTG 4385

Best Local Similarity 75.4%; Pred. No. 0; Matches 7326; Conservative 0; Mismatches 2267; Indels 125; Gaps 4;

QY	6	CCCTTAATAGGGCGACACTCCGCCATGAATCACTCCCTGTGAGGAATCTACTGTCTTCA	65
Db	7	CCCTGATGGGGCGACACTCCACCATGAATCACTCCCTGTGAGGAATCTACTGTCTTCA	66
QY	66	CGCAGAAAGCGTCTAGCCATAGGGGCTTAGTATAGTGTCTGATACAGCCTCCAGGCCCCCC	125
Db	67	CGCAGAAAGCGTCTAGCCATAGGGGCTTAGTATAGTGTCTGATACAGCCTCCAGGCCCC	126
QY	126	TCCCGGAGAGCCATAGTGTCTGCGGAACCGGTGAGTACACCGGAATTCGCGGAAGAC	185
Db	127	TCCCGGAGAGCCATAGTGTCTGCGGAACCGGTGAGTACACCGGAATTCGCGGAAGAC	186
QY	186	TGGGTCTCTTCTTGGATAAACCCACTCTATGCCCGGCCATTTGGGGTGGCCCGCAAG	245
Db	187	CGGGTCTCTTCTTGGATAAACCCCGCTCAATGCTGGAGATTTGGGGTGGCCCGCAAG	246
QY	246	CTGCTAGCCGAGTAGGCTTGGGTTGCGAAGGCTTTGTGTACTGCTGATAGGGTCTT	305
Db	247	CTGCTAGCCGAGTAGGCTTGGGTTGCGAAGGCTTTGTGTACTGCTGATAGGGTCTT	306
QY	306	GGAGTGCCCGGGAGGTCTCTGATAGCCGTTGCAACCGTGCACCATGAACAAATCTTAA	365
Db	307	GGAGTGCCCGGGAGGTCTCTGATAGCCGTTGCAACCGTGCACCATGAACAAATCTTAA	366
QY	366	GAABACCAAGAAACCAACCGTGCACCAAGACGTTAAGTTTCCGGGCGGGCC	425
Db	367	GAABACCAAGAAACCAACCGTGCACCAAGACGTTAAGTTTCCGGGCGGGCC	426
QY	426	AGATCGTTGGCGAGTATCTCTGTCGCGCAGGCGGCCAGGTTGGGTGTCGCGCGCA	485
Db	427	AGATCGTTGGCGAGTATCTCTGTCGCGCAGGCGGCCAGGTTGGGTGTCGCGCGCA	486
QY	486	CAAGGAAGACTTCGAGAGCGTCCCGACCGTGGGAAGCGCCAGCCCATCCCTAAAGATC	545
Db	487	CAAGGAAGACTTCGAGAGCGTCCCGACCGTGGGAAGCGCCAGCCCATCCCTAAAGATC	546
QY	546	GGCGCTCCACTGCGAAATCTTGGGGAACACGAGTATCCCTGGCCCTTATACGGGAATG	605
Db	547	GGCGCTCCACTGCGAAATCTTGGGGAACACGAGTATCCCTGGCCCTTATACGGGAATG	606
QY	606	AGGACTTCGGCTGGCAGAGTGTCTCTGTCGCCCGGAGGTTCCCGTCTCTTGGGGCC	665
Db	607	AGGACTTCGGCTGGCAGAGTGTCTCTGTCGCCCGGAGGTTCCCGTCTCTTGGGGCC	666
QY	666	CCAATGACCCCGGCATAGGTCCGCAACGTTGGGTAAGGTTCATCGATACCCCTAAAGTGG	725
Db	667	CCAATGACCCCGGCATAGGTCCGCAACGTTGGGTAAGGTTCATCGATACCCCTAAAGTGG	726
QY	726	GCTTTGCGGACCTCATGGGGTACATCCCTGTGTTGGGGCGCCGCTCGGGGGGTGCGCA	785
Db	727	GCTTTGCGGACCTCATGGGGTACATCCCTGTGTTGGGGCGCCGCTCGGGGGGTGCGCA	786
QY	786	GAGCTCTCGCGCATGCGTGAGAGTCTCGGAGACGGGGTTAATTTTGAACAGGGAAT	845
Db	787	GAGCTCTCGCGCATGCGTGAGAGTCTCGGAGACGGGGTTAATTTTGAACAGGGAAT	846
QY	846	TACCCGGTGTCTCTTTTCTATCTTCTGTCGCGCCCTGCTGCTCATCACACCCCGG	905
Db	847	TACCCGGTGTCTCTTTTCTATCTTCTGTCGCGCCCTGCTGCTCATCACACCCCGG	906
QY	906	TCTCCGCTCGGAGTGAAGAACATCAGTACCGGCTACATGTTGACTAACGACTGACCA	965
Db	907	TCTCCGCTCGGAGTGAAGAACATCAGTACCGGCTACATGTTGACTAACGACTGACCA	966
QY	966	ATGACAGCATTAACCTGCGAGCTCCAGGCTGCTGCTCCAGGTCGCCGGGTGCTCCCGT	1025
Db	967	ATGACAGCATTAACCTGCGAGCTCCAGGCTGCTGCTCCAGGTCGCCGGGTGCTCCCGT	1026
QY	1026	GCAGAAAGTGGGGAATGCACTCTCAGTGTGGATACCGGTCTCACCGAATGTGGCGTGC	1085

Db	1027	GCAGAGAAATGGGGAAATGCATCTCAGTGTGGATACCGGTCTCACCGAATGTGGCGTGC	1086
QY	1086	AGCGGCCCGGCCCTCAGCGAGGCTTCGCGAGGCACACATCGACATGTTGTGATCTCG	1145
Db	1087	AGCGGCCCGGCCCTCAGCGAGGCTTCGCGAGGCACACATCGACATGTTGTGATCTCG	1146
QY	1146	CCACGCTCTGCTCTGCCCTCTACGTTGGGGACCTCTGCGGTGGGTGATGCTCGACGCC	1205
Db	1147	CCACGCTCTGCTCTGCCCTCTACGTTGGGGACCTCTGCGGTGGGTGATGCTCGACGCC	1206
QY	1206	AAATGTTCTGCTCTGCCCGGACGACCACTGTTTGTCAAAGCTGCAATGTTGCTCATCT	1265
Db	1207	AAATGTTCTGCTCTGCCCGGACGACCACTGTTTGTCAAAGCTGCAATGTTGCTCATCT	1266
QY	1266	ACCTGGTACCATCACTGGACACCGCATGGATGGGACATGATGAATGTTGCTCGCCCA	1325
Db	1267	ACCTGGTACCATCACTGGACACCGCATGGGATGGGACATGATGAATGTTGCTCGCCCA	1326
QY	1326	CGGCTACCATGATCTTGGCGTACGCGATGCTGCCGAGGTCAATATAGACATCATTTA	1385
Db	1327	CGGCTACCATGATCTTGGCGTACGCGATGCTGCCGAGGTCAATATAGACATCATTTA	1386
QY	1386	GGGGGCTCATTTGGGGCGTCACTGTTGCGCTTGCGCTTCTCTATGACAGGAGGTGG	1445
Db	1387	GGGGGCTCATTTGGGGCGTCACTGTTGCGCTTGCGCTTCTCTATGACAGGAGGTGG	1446
QY	1446	CGAAGTCTGTTGCTCATCTTCTGTTGGCGCGGGGTGGACGCGCATCATCTGTTG	1505
Db	1447	CGAAGTCTGTTGCTCATCTTCTGTTGGCGCGGGGTGGACGCGCATCATCTGTTG	1506
QY	1506	GGGGTCTTCCCGCGCAGACCAACCGGCGCTCACCAAGCTTATTTGACATGGGCCCCAGGC	1565
Db	1507	GGGGTCTTCCCGCGCAGACCAACCGGCGCTCACCAAGCTTATTTGACATGGGCCCCAGGC	1566
QY	1566	AGAAATCCAGTCTGTTAAACCAATGAGGTGGGACATCAACCGCACCGCCCTGAACT	1625
Db	1567	AGAAATCCAGTCTGTTAAACCAATGAGGTGGGACATCAACCGCACCGCCCTGAACT	1626
QY	1626	GCAATGACTCTTGCACACCGGCTTATGCGGTCTCTGTTCTACACCCACAGCTTCACT	1685
Db	1627	GCAATGACTCTTGCACACCGGCTTATGCGGTCTCTGTTCTACACCCACAGCTTCACT	1686
QY	1686	CGTCAGGATGTCGCCGAGCATGTCGCTGCGCAGTATCGAGGCTTTCGGGTGGGAT	1745
Db	1687	CGTCAGGATGTCGCCGAGCATGTCGCTGCGCAGTATCGAGGCTTTCGGGTGGGAT	1746
QY	1746	GGGGCGCTTGAATATGAGGATAATGCAACCAATCCAGAGATATGAGACCTTATGCT	1805
Db	1747	GGGGCGCTTGAATATGAGGATAATGCAACCAATCCAGAGATATGAGACCTTATGCT	1806
QY	1806	GGCACTACCCCAAGGAGTGTGGGTGCTTCCGCGAAGACTGTGTGGGCCAGTGT	1865
Db	1807	GGCACTACCCCAAGGAGTGTGGGTGCTTCCGCGAAGACTGTGTGGGCCAGTGT	1866
QY	1866	ACTGTTTCAACCCAGCGAGTGTGGGTGCTTCCGCGAAGACTGTGTGGGCCAGTGT	1925
Db	1867	ACTGTTTCAACCCAGCGAGTGTGGGTGCTTCCGCGAAGACTGTGTGGGCCAGTGT	1926
QY	1926	ACACTGGGGGAGAAATGAGACAGATGTTCTTCTAATGAACAGCACTCGACCGCTGG	1985
Db	1927	ACACTGGGGGAGAAATGAGACAGATGTTCTTCTAATGAACAGCACTCGACCGCTGG	1986
QY	1986	GCTCATGTTGGGCTGACAGTGAATCTTCTGCTTACACCAAGACTTTCGGGGCAC	2045
Db	1987	GCTCATGTTGGGCTGACAGTGAATCTTCTGCTTACACCAAGACTTTCGGGGCAC	2046
QY	2046	CACCTGCGCTACTAGAGCTGACTTCAACCGCAGACCGGACCTGTTGTCGCCCAAGACT	2105
Db	2047	CACCTGCGCTACTAGAGCTGACTTCAACCGCAGACCGGACCTGTTGTCGCCCAAGACT	2106
QY	2106	GTTTAGGAAGCATCTGATACCATTAACCTCAATGCGGCTCTGGGCCCTGGCTCACGC	2165
Db	2107	GTTTAGGAAGCATCTGATACCATTAACCTCAATGCGGCTCTGGGCCCTGGCTCACGC	2166

Copied from 09540843 on 05/19/2004

9303	GATTATACCGGGTGGTTACCGTGGCGCGCGGGGGCGACATTTATCAGACGCTGTCCG	9366
9362	CGTCCGACCCCGCTATTATGCTCTTTTGGCCCTACTCTCTACTTTTTTGTAGGGGTAGGCCTT	9421
9363	CATGCGACCCCGCTATTACTCTCTTTGGCCCTACTCTCTACTTTAGCGTAGGAGTAGGCGATC	9422
9422	TTCTACTCCCGCTCGGTAGAGCGGCACACATTTAGCTTACACTCCATAGCTAACTGTCCC	9481
9423	TTTTTACTCCCGCTCGGTAGAGCGGCAACACCTAGCTTACACTCCATAGCTTAGTTCCGT	9482
9482	TT	9510
9483	TT	9511

RESULT 7
 AAC86645
 AAC86645 standard; DNA; 9611 BP.
 AAC86645;
 02-APR-2001 (first entry)
 Nucleotide sequence of chimeric Hepatitis C virus clone pH77CV-J6S.
 HCV; HCV strain HC-J6CH; HCV genotype 2a; antiviral; vaccine; ss.
 Synthetic.
 Hepatitis C virus.
 Key Location/Qualifiers
 CDS 342..9389
 /*tag= a
 WO200075338-A2.
 14-DEC-2000.
 02-JUN-2000; 2000WO-US15446.
 04-JUN-1999; 99US-0137693.
 (USSH) US DEPT HEALTH & HUMAN SERVICES.
 Yanagi M, Bukh J, Emerson SU, Purcell RH;
 WPI; 2001-061728/07.
 Nucleic acid molecule encoding human hepatitis C virus of genotype 2a for developing vaccines, for diagnosis of hepatitis C virus and in screening assays for identification of antiviral agents -
 Disclosure; Page 100-103; 167pp; English.
 AAC86645-48 represent chimeric cDNA clones of infectious Hepatitis C virus (HCV). In each clone the C, E1 and E2 genes are derived from HCV strain HC-J6CH genotype 2a. The p7 protein was derived from HCV strain HC-J6CH or HCV strain PCV-H77C, and the NS genes were derived from HCV strain PCV-H77C. Such HCV sequences are capable of expressing the virus when transfected into cells. The HCV protein is useful for assaying candidate antiviral agents for activity against HCV. Antibodies specific for HCV polypeptide are useful in prevention and treatment of diseases caused by HCV in animals, in particular humans. The HCV polypeptides serve as immunogens in the development of vaccines for preventing HCV in mammals or as antigens in diagnostic assays for detecting the presence of HCV in biological samples. The HCV polynucleotide is also useful for identifying cell lines capable of supporting the replication of HCV in vitro and to produce attenuated viral strains via passage in vitro or in vivo.
 Sequence 9611 BP; 1935 A; 2896 C; 2682 G; 2098 T; 0 other;
 Query Match 59.7%; Score 5800.8; DB 22; Length 9611;

5923 GGGGCAAGCCTGGTAGCAGACTTACACCTGCTGAGACTACGGTGAAGCTCCGGGCTATTTC 4988

5982 AACACGCCCGGTTTGCCTGTGTGCAAGAACCACTCTGAGTTTTTGGAGGCAAGTTTTTCACC 5041

5983 AACACGCCCGGTTTGCSCGTATGTCAAGACCACCTGGAGTTCTGGGAAGCGGTCTTTACA 5042

5042 GGGCTCACACACATAGATGCCCACTTCTCTTCCAAACAAGCAATCGGGGGAATTTTC 5101

5043 GGTCTCAWCACATTTACCGCCCACTTCTCTCCAGACGAAAGAGGAGAGAAACTTT 5102

5102 GCATCTTAAACAGCCTACACAGCTACAGTGTGGCTAGGCGCAAGAGCCGCCCCCGTCC 5161

5103 GKTATCTAACGGCCCTACAGGCCACAGTATGCGCCAGGGCAAGGCCCTCTCTCTTGG 5162

5162 TGGGAGCTCATGTGGAAGTGTGTGACTCGACTCAAGGCCCACTCGTGGGCCCCACACCT 5221

5163 TGGGAGCTGATGTGGAAGTGTCTAACTAGGCTCAAACTACACTGACTGGTCCCACCC 5222

5222 CTCCTGTACCGTCTGGGCTCTGTTACCAACGAGTCAACCTTCAACATCCCGTGACGAAA 5281

5223 CTCCTGTACCGTCTGGGTCCGCTGAGCAATAGGTACCTTGAACGACACCCCGGTGACGAAA 5282

5282 TACATCGCCACCTGCATGCAAGCGGACCTTCAGGTCAATGACCAGCACTAGGGTCTTGGCA 5341

5283 TACATCGCCACCTGCATGCAAGCTGACCTYAGATCATGACAGCTCATGGGTCTTGGCG 5342

5342 GGGGAGTCTTGGCGGCCGTGCGCGGTATTGCCCTGGCGACCGGGTGTGTTGCATCATC 5401

5343 GGGGGGTCTAGCGCGCGTGGCAGCTTACTGCTGGCGACTGGCTGCAATTTCCATCAT 5402

5402 GGGCGTCTGCACATTAACAGGAGCGGTCTGTTGGCGGGAACAGGAGTCTCTATGAG 5461

5403 GGGCGCTTACACTCAATGATCGGGTGTGTGTRCCCYCAGACAGGARATCTTATATGAG 5462

5462 GCTTTGTAGAGATGGAGGATGTGCCCTTAGGCGGCTCTCATTTGAAGAGGGCGCAGCG 5521

5463 GCCTTTGTAGAGATGGAGGATGTGCCCTTCAAAGCGCCCTCATTTAGGAAGAGGCGCAGCG 5522

5522 ATAGCCGAGATGTGAAGTCCAAGATCCAAGGCTTATTGACAGCAAGCTTCCAACCAAGCT 5581

5523 ATGGCGGAGTGTCAATCTAAGATACAAGGCTCTCTACAACAGGCCACAAGGCAAGCT 5582

5582 CAAGACATACACCCACTGTGACGCTTCATGGGCCAAGGTAGAACAAATTTCTGGGCCAAA 5641

5583 CAAGCATRCAGCAGCTATACAGTCATCATGGGCCAAGCTTGAAACAAATTTTGGGCCAAA 5642

5642 CACATGTGGAATTTCACTTAGCGGCATCAATACCTCGCAGGACTATCAACACTGSCCAGGG 5701

5643 CACATGTGGAATTTCACTAGTGGTATACAGTACCTTAGCAGGACTCTCAACCTACCGGA 5702

5702 AACCTGCAGTAGCTTCCAATGATGGCGTTTCAGTGGCGGCTTACCAGTCCGCTGTCAACA 5761

5703 ANTCTGCAAGTGCATCAATGATGCTTTAGCGCGCGCTGACTAGCCCACTACCCACC 5762

5762 AGCACCATATCTTCTCAACAAATTTTGGGGGCTGGCTAGCATCCGAAATTTGCACACACC 5821

5763 AGCACCATCTCTTTGAACATCATGGGAGGATGCTTGGGCTCTYCAAGTTGGCCCCCT 5822

5822 GCGGGGCCCATCTGGCTCTGCTGCTAGTGGGCTAGTGGGAGCTGCCGTAGGCAAGTATAGGC 5881

5823 GCGGAGGACCATYGGCTCTGTTGTAGTGGTCTAGTGGGGGCGCGCTCGGAGACATAGGC 5882

5882 TTAGGTAAGTGTCTAGTGGACATCTCTGGCAGGATATGTTGGCGGCAATTTGGGGGCTCTC 5941

5883 CTGGGTAAGNATCTGTGTGACGTTTTTGGCCGGTACGGCGAGGCAATTTCAAGGGGCCCTC 5942

5942 GTGCGCATTCAGATCATGTCTGGCAGAGAGCCCTCCATGGAGGATGTGCTCAACTTGGCTG 6001

5943 GTAGCTTTTAAGATCATGAGCGCGAGAGCCACGGTAGAAGACGTTGTGAATCTCTCTG 6002

6002 CTTGGAAATCTGTCTCGGGTCTGTTGGTAGTGGGAGTCACTTGGCGGGCAATTTCTGGC 6061

6003 CCTGCTATYCTGTCTCTGGTGTGGTGTGGAGTCACTCTGTGACGAAATTTCTGGC 6062

1682 AACTCGTCAGGATGTCCTCCGAAACGATGTCGCTCGCCAGTATCGAGGCTTCCGGGTG 1741
1683 AACAGCTCTGGCTGCCCCGAGGCTTGTCTTCCTGCGCGGGCTGGACGATTTTCGATC 1742
1742 GGATGGGGCCCTTGCATATGAGGATATAGTACCAATCCAGAGATATGAGACCTTAT 1801
1743 GGGTGGGAACTTGGAAATACGAAACCAACGTCACCAACGATGAGACATGAGCCGTAC 1802
1802 TGCTGGCACTACCCACCAAGGAGTGTGGGCTGCTCCGCGAAGACTGTGTGTGCCCA 1861
1803 TGCTGGCACTACCCACCAAGGAGTGTGGGCTGCTCCGCGAAGACTGTGTGTGCCCA 1862
1862 GTGTACTGTTTACCCCGAGGCTTGTGTGGGCAACGACCAAGGCTTGTGGGCGCC 1921
1863 GTCTATGTTTACCCCGAGGCTTGTGTGGGCAACGACCAAGGCTTGTGGGCGCC 1922
1922 ACTTACAGTGGGGGAGATAGACAGATGTCCTTCTTATTTGAACAGCACTCGACACCG 1981
1923 ACCTACACCTGGGGGAGAAACGAGACCGATGTCCTTCTGCTRAATAGCAACAGACCCCG 1982
1982 CTGGGCTCAGTGTTCGGCTGCACTGGATGAACTCTTCTGGCTACACCAAGACTTGGGC 2041
1983 CGAGGAGCTGGTTCGGCTGCACTGGATGAACTGGGCTTCTAAGACATGCGGT 2042
2042 GCACCACTTCCGCTACTAGAGCTGACTTCAAGCCGAGACGAGCCTGTTGTGCCCAAG 2101
2043 GCACCACTTCCGCTACTAGAGCTGACTTCAAGCCGAGACGAGCCTGTTGTGCCCAAG 2102
2102 GACTGTTTGAAGAGATCTGTATACACTTCACTCAATTCGAGGCTGTCGGGCTGGCTC 2161
2103 GACTGTTTGAAGAGATCTGTATACACTTCACTCAATTCGAGGCTGTCGGGCTGGCTC 2162
2162 ACCCGAGTGTCTGATGACTACCTTACGCTCTGAGCTTACCTTACCTGCAAGTTAAC 2221
2163 ACTCCAGGCTCTGAGTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTT 2222
2222 TATACCACTTCAAAATAGGATGATGTGGAGGAGGTTGAGCAGAGGCTCAGCGTGCA 2281
2223 TACCACTTCAAAATAGGATGATGTGGAGGAGGTTGAGCAGAGGCTCAGCGTGCA 2282
2282 TGCAATTCATCTGCGGAGATCTGCACTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2341
2283 TGCAATTCATCTGCGGAGATCTGCACTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2342
2342 CTTTGTGACTTCCACGAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2401
2343 CCACTGCTGATTCACACTGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2402
2402 GCCTTGTGAGTGGTCTTCCACCTCCACCAACATCGTGGAGGAGGAGGAGGAGGAGGAGG 2461
2403 GCATATCCACTGGCTTATGCACTTCCACCAACATCGTGGAGGAGGAGGAGGAGGAGGAGG 2462
2462 GGCCTATCACTGCTTCCACCAACATCGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2521
2463 GGCCTATCACTGCTTCCACCAACATCGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2522
2522 CTGCTCTTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2581
2523 TTGTTGTGGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2582
2582 GCGAAGCAGCAGTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2641
2583 GCGAAGCAGGAGTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2642
2642 GGCCTATCACTGCTTCCACCAACATCGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2701
2643 GGCCTATCACTGCTTCCACCAACATCGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2702
2702 CCCTTAGCTACCTTATCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2761
2703 CCGTGGCAGCAGTACTGTGCTCGGCTTGTGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 2762

2762 CCCAACAGGCTTATGCTTATGACGATCTGTGCAATGCGCAGATAGGAGCGCTCTGTGCTG 2821
2763 CCACAGGAGCTTATGCTTGTGAGCCTGTGAAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2822
2822 GTAATGATCACTCTCTTACTCTACCCCGGGGTATAGACCTTCTCTAGCCGGTCTTTTGG 2881
2823 GTAATGATCACTCTCTTACTCTACCCCGGGGTATAGACCTTCTCTAGCCGGTCTTTTGG 2882
2882 TGGTGTGTGTGTCTTCTCTGACCTGCGGGGAGGATGATGCTCCAGGAGTGGGACCACT 2941
2883 TGGTGTGTGTGTCTTCTCTGACCTGCGGGGAGGATGATGCTCCAGGAGTGGGACCACT 2942
2942 ATGCAAGGCTGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3001
2943 CTGAGGAGTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3002
3002 GTGGTGTGTGATCACTTAAACAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3061
3003 CTGTTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3062
3062 GGTGTGTGTGACGCGGCTGCGTACTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3121
3063 GGTGTGTGTGACGCGGCTGCGTACTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3122
3122 ATGCAAGGCTGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3181
3123 CTGTTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3182
3182 TGAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3241
3183 TGAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3242
3242 CTGCGGAG 3301
3243 TTTTGGGAG 3302
3302 ATTTGTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3361
3303 ATTTGTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3362
3362 GCGGAGCTTGTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3421
3363 GCGGAGCTTGTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3422
3422 AGTCTTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3481
3423 AAKTCTTGTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3482
3482 GTGTGTGAG 3541
3483 GTGTGTGAG 3542
3542 ACAGTCTGAG 3601
3543 TCGGTCACACAACTTCTTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3602
3602 GGAGCTGCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3661
3603 GGGGTGTGTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3662
3662 GCTGAG 3721
3663 GCAGAG 3722
3722 AGTGTGAG 3781
3723 ACCTGCGGAG 3782
3782 AGAGCGGAG 3841
3783 AGGAAGAG 3842
3842 TCCTCAGGAG 3901

XX PF 30-JUL-1992; 92EP-0306952.
 XX PR 09-AUG-1991; 91JP-0287402.
 XX PR 05-DEC-1991; 91JP-0360441.
 XX PA (IMMO) IMMUNO JAPAN INC.
 XX PI Nakamura T, Okamoto H;
 XX DR WPI; 1993-087166/11.
 XX DR N-PSDB; AAR33539, AAR33214.
 XX PT Polynucleotide(s), polypeptide(s) and antibodies of NANBH virus -
 XX PT useful for detecting NANBH, as a vaccine and for screening blood
 XX PT samples
 XX PS Claim 7; Page 59-64; 93pp; English.
 CC RNA was isolated from the plasma of human patients positive for
 CC NANBH virus (strain HC-J8) and was subjected to reverse transcription
 CC to produce cDNA. The resulting cDNA was amplified by PCR, and
 CC nucleic acid sequences determined by analysis of clones obtd. by PCR
 CC amplification (42 clones in total). The NANBH HC-J8 degenerate genome
 CC was found to contain a degenerate open reading frame encoding
 CC polypeptide precursors of 303 amino acid residues.
 CC See also AA038172-220.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX SQ Sequence 9511 BP; 2020 A; 2709 C; 2539 G; 2108 T; 135 other;
 Query Match 62.1%; Score 6029.8; DB 14; Length 9511;
 Best Local Similarity 76.3%; Pred. No. 0;
 Matches 7256; Conservative 105; Mismatches 2148; Indels 0; Gaps 0;
 2 CCGGCCCTAATAGGGCGACACTCCGCCATGAATCACTCCCTGTGAGGAATCTACTGTC 61
 3 CCGGCCCTCATAGGGCGGACACTCCGCCATGAATCACTCCCTGTGAGGAATCTACTGTC 62
 62 TTACGCGAAGAGCGTACCGATGGCGTGTAGTATGATGCTGTACAGCGCTCCAGGCC 121
 63 TTACGCGAAGAGCGTCTAGCGATGGCGTGTAGTATGATGCTGTACAGCGCTCCAGGCC 122
 122 CCCCTCCGGGAGAGCCATAGTGTCTCGCGAACCCTGTAGTACACCGGAATTCGCGGA 181
 123 CCCCTCCGGGAGAGCCATAGTGTCTCGCGAACCCTGTAGTACACCGGAATTCGCGGA 182
 182 AGACTGGGTCTTTCTTGGATAAACCACTCTATGCCCCGCCATTTGGCGGTGCCCCCGC 241
 183 AGACTGGGTCTTTCTTGGATAAACCACTCTATGCCCCGCCATTTGGCGAAGCCCCCGC 242
 242 AAGACTGCTAGCCGAGTAGGTGGGTTCGGAAGGCCCTTGTGTACTGCTGTAGGT 301
 243 AAGACTGCTAGCCGAGTAGGTGGGTTCGGAAGGCCCTTGTGTACTGCTGTAGGT 302
 302 GCTTGCAGTGGCCCCGGAGGTCTGTAGACCGGTGACCATGAGCACAATCTCTAAACCT 361
 303 RCTTGCAGTGGCCCCGGAGGTCTGTAGACCGGTGACCATGAGCACAATCTCTAAACCT 362
 362 CAAGAAGAAACCAAGAAACACCAACCGTGCGCCCAACAAGCTTAAGTTTCGGGGCGC 421
 363 CAAGAAGAAACCAAGAAACACCAACCGTGCGCCCAACAAGCTTAAGTTTCGGGGTGC 422
 422 GGCCAGATCTGTGGCGAGTATATCTGTGCGCGCAGAGGCCGCCAGGTGGGTGTCGC 481
 423 GGTGAGATCTGTGGCGAGTATATCTGTGCGCGCAGAGGCCGCCAGGTGGGTGTCGC 482
 482 GCGACAAGAGAGACTTCGAGCGGTCCCGACCGTGGAGAGGCCAGCCCATCCCTAAA 541
 483 GCGACAAGAGAGACTTCGAGCGGTCCCGACCGTGGAGAGGCCAGCCCATCCCGAAA 542
 542 GATCGCGCTCCACTGGCAATCTTGGGGAACACAGGATACCCCTGGCCCCCTATACGG 601
 543 GATCGCGCTCCACTGGCAATCTTGGGGAACACAGGATATCCCTTGGCCCCCTGTACGA 602
 602 AATGAGGACTCGGCTGGGACAGGATGGCTCTGTCCTCCCGGAGGTTCCTGCTCTTGG 661
 603 AAGAGGGTGGGGCTGGGGTGGCTCTGTCCTCCCGGAGGTTCCTGCTCTTGG 662
 662 GGGCCCAATGACCCCGGCATAGGTGCGCAACGTTGGGTAAAGTTCATGATACCTTAAAC 721
 663 GGGCCCAACCGACCCCGGCATAGATCACGCAATTTGGGCGAGGTTCATGATACCTTAAAC 722
 722 TGGGGCTTGGCGACCTCATGGGTACATCCCTGCTGGGCGCCCGCTCCGCGCGCTC 781
 723 TGTGGTGGTGGCGACCTCATGGGTACATCCCTGCTGGGCGCCCGCTGGAGGCGTC 782
 782 GCCAGAGCTCTCGCGCATGGCGTGTAGGTCTCTGGAGACGGGTAAATTTGCAACAGGG 841
 783 GCCAGAGCTCTGGCACACGGTGTAGGTCTCTGGAGACGGGTAAATTTGCAACAGGG 842
 842 AACTTACCGGTGTCTCTTTTCTATCTTCTTGTGGCGCTGTGTCTGTGATCACCAACC 901
 843 AATTTACCGGTGTCTCTTTTCTATCTTCTTGTGTGTCTGTGTGTGTGTGTGTGTGT 902
 902 CCGGTCTCGCTCGCGAAGTGAAGACATCAGTACCGGCTACATGCTGTGTGTGTGTGTGT 961
 903 CCGGTCTCTGCGATGGAGTGTAGGTCTGTAGTGTGTGTGTGTGTGTGTGTGTGTGT 962
 962 ACCAATGACAGCATTACCTGGCGCTCCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1021
 963 TCAACACAGCATCCTCTGGCGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1022
 1022 CCGTTCGAGAAAGTGGGAAATGATCTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1081
 1023 CCATGTGAGAAATGATGAGGACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1082
 1082 GTGCGAGCGCGCGCGCTCTCGAGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1141
 1083 GTGAAACACCGCGGT 1142
 1142 TCGCGCACGTCTGTCTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1201
 1143 GCGACTGTGCGCTGT 1202
 1202 GCGCAATGTTCATGTCTGT 1261
 1203 TCGAGGCTTTCATGTATCACCAACCGCACTTTCACCCAGAGTGTGTGTGTGTGTGT 1262
 1262 ATCTACCTGT 1321
 1263 ATCTACCAAGT 1322
 1322 CCACGCTTACCATGT 1381
 1323 CCACCTTTCATGT 1382
 1382 ATTAGCGGGCTCATTTGGGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1441
 1383 ATTTTCGGCGCATTTGGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1442
 1442 TGGGCAAGAGT 1501
 1443 TGGGCAAGAGT 1502
 1502 GTTGGGGT 1561
 1503 ASGGT/CAGSAGCGGT 1562
 1562 AGGCAAGAAATCAGCTGT 1621
 1563 AAGCAGAACTCTATTTTATCAACCAATGGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1622
 1622 AACTGCAATGATCTGT 1681
 1623 AATTGCAATGACAGCTTASAGAGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1682

```

Db      7861  |||||CTCAGTCTTAAAGACATCAAGCTAGCGGCTTCAAGGTCAAGGCTCCTCACCTT 7920
Qy      7921  |||||GGAGGAGCTTCCAGTTAAACCCACCCCACTTCTGCAAGATCTAAATATGGTTTGGGGC 7980
Db      7921  |||||GGAGGAGGCTGCGAGTTGATCTCCACCCCACTTCTGCAAGATCCAAAGTATGGATTCGGGGC 7980
Qy      7981  |||||TAAGGAGTCCGAGCTTGTCCGGAGGCGGTTAAACCAATCAATCAAGTCCGTGTGGAAGGA 8040
Db      7981  |||||CAAGGAGTCCGAGCTTGTCCGGAGGCGGTTAAACCAATCAAGTCCGTGTGGAAGGA 8040
Qy      8041  |||||CCTCTGAGGAGTCTAGAAACACCAATTTCCCAACCAATTTATGCGCAAAATAGAGTGT 8100
Db      8041  |||||CCTCTGGAAGACCCCAACCAACCAATTTCCCAACCAATTTATGCGCAAAATAGAGTGT 8100
Qy      8101  |||||CTGCGTGGACCCCAACCAAGGGGGGCAAGAGAGCTCGCTTATGTTTACCTGACCT 8160
Db      8101  |||||CTGCGTGGACCCCAAGGGGGGTAAGAAACAGCTCGCTCATCGTTTACCTGACCT 8160
Qy      8161  |||||CGCGCTCAGGCTCTGCGAGAGATGGCCCTTTATGACATTTACACAAACCTTCTCAGGC 8220
Db      8161  |||||CGCGCTCCGGGTCTGCGAGAAATGGCCCTCTATGACATTTACAAAGCTTCTCAGGC 8220
Qy      8221  |||||GGTGATGGGGCTTCTTATGATTTCCAGTATTTCCCGCTCAGCGGTAGAGTTTCTTT 8280
Db      8221  |||||GGTAAATGGGAGCTTCTTATGGCTTCCAGTACTCCCTCGCCCAACGGGTGAGTATCTTT 8280
Qy      8281  |||||GAAAGCATGGCGGAAAGAGAACCTATGGGTTTTTGTATGATACCGGATGCTTTGA 8340
Db      8281  |||||GAAAGCATGGCGGAAAGAGAACCTATGGGTTTTTGTATGATACCGGATGCTTTGA 8340
Qy      8341  |||||CTCAACCGTCACTGAGAGACATCAGGAGTCCATATATCGGCGCTGCTCCTT 8400
Db      8341  |||||CTCAACCGTCACTGAGAGACATCAGGAGTCCATATATCGGCGCTGCTCCTT 8400
Qy      8401  |||||GCCCGAGGAGGCGGCACTGCCATACACTCGCTGACTGAGAGACTTTACGTAGGAGGCC 8460
Db      8401  |||||GCCCGAGGAGGCGGCACTGCCATACACTCGCTGACTGAGAGACTTTACGTAGGAGGCC 8460
Qy      8461  |||||TATGTTCAACAGAGGCGCAACCTCGGGTACAGCGGTGCGCGCCAGCGGGTGCT 8520
Db      8461  |||||CATGTTTCAACAGAGGCGTCAACCTCGGGTTACAGACGTTGCGCGCCAGCGGGTGCT 8520
Qy      8521  |||||CACCCTAGCATGGGAAACACCATCATCTGCTAGGTGAAGCCTTAGCGGCTTGTAAAGC 8580
Db      8521  |||||AACCCTAGCATGGGTAAACCATCATCTGCTATGTGAAGCCTTAGCGGCTTGTAAAGC 8580
Qy      8581  |||||TGCGGGATTAATCGCGCCCACAATGCTGGTATGCGCGGATGACTTGGTTGTCTCTAGA 8640
Db      8581  |||||TGCGGGATTAATGCGCGCCCACAATGCTGGTATGCGCGGATGACTTGGTTGTCTCTAGA 8640
Qy      8641  |||||AAGCCAGGAGACCGAGAGGACGAGGGAACCTGAGAGCCTTACCGAGGCTATGACCAAG 8700
Db      8641  |||||AAGCCAGGAGACCGAGAGGACGAGGGAACCTGAGAGCCTTACCGAGGCTATGACCAAG 8700
Qy      8701  |||||GTATTTGCGCCCTCTGCTGACCCCGCCAGCGGAGTATGATCTGAGCTGATACATC 8760
Db      8701  |||||GTACTTGGCCCTCTCTGCTGATCCCGGACCGGAATGACTGGAGCTAATACATC 8760
Qy      8761  |||||TTGCTCTCAATATGCTCTGCTGGCGTGGGCGCCCAAGGCGCGGAGATCTTACCTGAC 8820
Db      8761  |||||CTGTCTCTCAATATGCTCTGCTGGCGTGGGCGCGGCGCGGAGATCTTACCTGAC 8820
Qy      8821  |||||CAGAGACCTTACCATCTCCAAATCGCGGGCTGCTGGGAAACAGTTAGACATCTCCCTGT 8880
Db      8821  |||||CAGAGACCTTACCATCTCCAAATCGCGGGCTGCTGGGAAACAGTTAGACATCTCCCTGT 8880
Qy      8881  |||||CAATTCATGCTGGGAAACATCATCCAGTACCGCCGACCATATGGCTCGATGCTCT 8940
Db      8881  |||||CAATTCATGCTGGGAAACATCATCCAGTACCGCCGACCATATGGCTCGATGCTCT 8940
Qy      8941  |||||GATGACACATCTTCTTCTCAATCTCTATGGCTCAAGACAGCTGGAGCAGAACCTCAACTT 9000

```

```

Db      8941  |||||AATGACACACATTTCTTCTCCATTTCTCATGGTCCAGACACCTTGGACAGAACCTCAACTT 9000
Qy      9001  |||||TCAGATGTACGAGCGGTGTACTCCGTGAGTCCCTTGGACCTCCCGAGCTATAATTCAGAG 9060
Db      9001  |||||TCAGATGTATGGATCAGTATATCTCCGTGAATCTCTTGGACCTTCCAGCCATATTCAGAG 9060
Qy      9061  |||||GTTATACATGGGCTTACCGCTTTTCTCTGCACACATACCTCCCGACCAATCTACACCGGT 9120
Db      9061  |||||GTTACACGGCTTACCGCTTTTCTCTATGCACACATATCTCTACACCAATCTACACCGGT 9120
Qy      9121  |||||GGCTTACGCTTCAGAAACATTTGGGGCGCACCTCTCAGAGCGTGGAGAGCGGGGACG 9180
Db      9121  |||||GGCTTACGCTTCAGAAACATTTGGGGCGCACCTCTCAGAGCGTGGAGAGCGGGGACG 9180
Qy      9181  |||||TCAGTTCAGGCGTCCCTCATCTCCGTGGGGGAGAGCGGCGTTCGGGTGATATCT 9240
Db      9181  |||||CGCAGTTCAGGCGTCCCTCATCTCCGTGGAGGAAAGCGGCGTTTGGCGCGATATCT 9240
Qy      9241  |||||CTTCAATTTGGGCGGTGAAGACCAAGCTCAAACTCACTTCCATTTGCCGGAAGCGGCTCT 9300
Db      9241  |||||CTTCAATTTGGGCGGTGAAGACCAAGCTCAAACTCACTTCCATTTGCCGGAAGCGGCTCT 9300
Qy      9301  |||||GGATTTATCCAGCTGGTTTCCAGTCCGCGCGCGCGGCGGAGACATTTATCAAGCGTGT 9360
Db      9301  |||||GGACTTATCCAGTTGGTTTCCAGTCCGCGCGCGCGGCGGAGACATTTATCAAGCGTGT 9360
Qy      9361  |||||GGTGGCGGACCCGCTTATTTGCTTTTGGGCTTACTCTTCTTTTGTAGGGGTAGGCT 9420
Db      9361  |||||GGCGCGCGGACCCGCTTATTTACTCTTCCGCGCTACTCTCTTCTTTTGTAGGGGTAGGCT 9420
Qy      9421  |||||TTTCTTACTTCCCGCTCGGTAGAGCGGACACATTAAGTACACTCCATAGTAACTGTCT 9480
Db      9421  |||||CTTCTTACTTCCCGCTCGGTAGAGCGGACACACTAGTACACTCCATAGTAACTGTCT 9480
Qy      9481  |||||CTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 9540
Db      9481  |||||CTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 9533
Qy      9541  |||||TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 9600
Db      9534  |||||TTTTTTTCCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 9567
Qy      9601  |||||TACTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 9660
Db      9568  |||||TACTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 9627
Qy      9661  |||||GAGCGCATGACTGCGAGAGTGGCGTAACTGGTCTCTCTGCGATCATGT 9711
Db      9628  |||||GAGCGCATGACTGCGAGAGTGGCGTAACTGGTCTCTCTGCGATCATGT 9678

```

RESULT 6

AAQ38221
ID AAQ38221 standard; cDNA to mRNA; 9511 BP.

XX AAQ38221;

XX AC AAQ38221;

DT 25-MAR-2003 (updated)

DT 01-JUL-1993 (first entry)

XX NANBH virus strain HC-J8 genome cDNA sequence.

XX Non A non B hepatitis virus; plasma; degenerate; ss.

XX Non A, non B hepatitis virus strain HC-J8.

XX Key Location/Qualifiers

FT CDS 342..943

FT /*tag= a

XX EP532167-A2.

XX 17-MAR-1993.

QY 5701 GAACCTGACAGTAGCTTCCATGATGGCTTTCAGTGCAGCCCTCACAGTCCGCTGTCAAC 5760
 DB |||||
 QY 5701 GAACCCCGCGGTCCATGATGGCAATTCAGTGCAGCCCTCACAGTCCGCTGTGCAC 5760
 DB |||||
 QY 5761 AAGCACATATCTTCTCAACATTTTGGGGGCTGCTAGCATCCCAAATTCACACACC 5820
 DB |||||
 QY 5761 CAGTACACCATCTTCTCAACATCATGGGAGCTGGTTAGCGTCCAGATCGCACACC 5820
 DB |||||
 QY 5821 CGCGGGGGCCACTGGCTTCTGTTGTCAGTGGCTAGTGGGAGCTCCGATAGGAGATAGG 5880
 DB |||||
 QY 5821 CGCGGGGGCCACCGGCTTCTGTCAGTGGCTGGTGGGGGCTGCGGTGGGAGCATAGG 5880
 DB |||||
 QY 5881 CTTAGGTAAAGTCTAGTGCACATCTTGGCAGGATATGGTGGGCAATTTCCGGGGCTCT 5940
 DB |||||
 QY 5881 CCTGGGTAAAGTCTGCTGACATCTTGGCAGGATATGGTGGGCAATTTCCGGGGCCCT 5940
 DB |||||
 QY 5941 CGTGCATTAAGATCATGTCTGGCGAAGAGCCCTCCATGAGAGATGCTCAATCTACT 6000
 DB |||||
 QY 5941 CGTGCATTAAGATCATGTCTGGCGAAGAGCCCTCCATGAGAGATGCTCAATCTACT 6000
 DB |||||
 QY 6001 GCCTGGAAATCTGCTCCGGGTGCTTGGTGGGAGTCAATCTGGCGGCAATCTGCG 6060
 DB |||||
 QY 6001 GCCTGGGATCTGCTCCGGGAGCCCTGGTGGGAGTCAATCTGGCGGCAATCTGCG 6060
 DB |||||
 QY 6061 CCGACAGTGGGACCGGGGGAAGCGCGCTCCATGATGAATAGACTCAATGCTTTGC 6120
 DB |||||
 QY 6061 CCGCAGTGGGACCGGGGGAAGCGCGCTCCATGATGAATAGACTCAATGCTTTGC 6120
 DB |||||
 QY 6121 TTCAGAGGAATCAAGTCTGCGGAGAGCCCTCCATGAGAGATGCTCAATGCTTTGC 6180
 DB |||||
 QY 6121 TTCAGAGGAATCAAGTCTGCGGAGAGCCCTCCATGAGAGATGCTCAATGCTTTGC 6180
 DB |||||
 QY 6181 TGTGACCAACTACTTGGCTCCCTTACCATAAACAGCTGCTCAGAGACTCCACACTG 6240
 DB |||||
 QY 6181 TGTGACCAACTACTTGGCTCCCTTACCATAAACAGCTGCTCAGAGACTCCACACTG 6240
 DB |||||
 QY 6241 GATTACTGAGACTGCCCTCCATCCAGCGGCTCGTGGCTCCGCGATGCTGGGACTG 6300
 DB |||||
 QY 6241 GATTAATGAGACTGCCCTCCATCCAGCGGCTCGTGGCTCCGCGATGCTGGGACTG 6300
 DB |||||
 QY 6301 GGTTCGACCATCTAAACAGACTTTAAATATGGCTGACTCCAAATTTATCCCAAAGAT 6360
 DB |||||
 QY 6301 GGTTCGACCATCTAAACAGACTTTAAATATGGCTGACTCCAAATTTATCCCAAAGAT 6360
 DB |||||
 QY 6361 GCCCGGCTCCCTTCT 6420
 DB |||||
 QY 6361 GCCCGGCTCCCTTCT 6420
 DB |||||
 QY 6421 CATCATGACACAGCTGCTCTGCGGCGCAATATCTTGGCAATGCTCGCTTGGGCTC 6480
 DB |||||
 QY 6421 CATCATGACACAGCTGCTCTGCGGCGCAATATCTTGGCAATGCTCGCTTGGGCTC 6480
 DB |||||
 QY 6481 CATGAGAAATCAAGGAGCTTAAAGCTTGAATATCTGCGAGGGGACCTTTCTATCAA 6540
 DB |||||
 QY 6481 TATGAGGATCACAGGGCTTAAACCTTGCATGAACACCTGGCAGGGGACCTTTCTATCAA 6540
 DB |||||
 QY 6541 TTGTTACACGAGGGGCTGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6600
 DB |||||
 QY 6541 TTGTTACACGAGGGGCTGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6600
 DB |||||
 QY 6601 GAGGGTGGGCGCTCAGAGTACGCGAGGAGTACGCGAGAGAGAGAGAGAGAGAGAGAGAG 6660
 DB |||||
 QY 6601 GAGGGTGGGCGCTCAGAGTACGCGAGGAGTACGCGAGAGAGAGAGAGAGAGAGAGAGAG 6660
 DB |||||
 QY 6661 AGGACTCACCAGTGAATCTTGAAGTCCCTGCGCAACTACCTCTCTCCCGAGTCTTTTC 6720
 DB |||||
 QY 6661 AGGACTCACCAGTGAATCTTGAAGTCCCTGCGCAACTACCTCTCTCCCGAGTCTTTTC 6720
 DB |||||
 QY 6721 CTGGGTGGAGCGAGTGCAGATCCATAGTTTGGCCCGGAGAGAGAGAGAGAGAGAGAGAG 6780
 DB |||||
 QY 6721 CTGGGTGGAGCGAGTGCAGATCCATAGTTTGGCCCGGAGAGAGAGAGAGAGAGAGAGAG 6780
 DB |||||

QY 6781 TGAGGTCTCTGTTCTGCGTTGGCTTAAATTAATTTGCTGCGGTCCAGCTTCTCTTGGCA 6840
 DB |||||
 QY 6781 TGAGGTCTCTGTTCTGCGTTGGCTTAAATTAATTTGCTGCGGTCCAGCTTCTCTTGGCA 6840
 DB |||||
 QY 6841 CCTTGAAACCGACACAGAGCTTATGATGTCATGCTAAAGATCCATCTCATATCAGGCG 6900
 DB |||||
 QY 6841 ACCTGAGCCCGACGACAGCTTATGAGGTCCATGCTAAAGATCCCGCCCGACATCAGGCG 6900
 DB |||||
 QY 6901 SGAGACTGACGCGCGGCTTTAGCGCGGGGTTCACCCCATCCGAGGCAAGCTCTCTCGG 6960
 DB |||||
 QY 6901 SGAGACTGACGCGCGGCTTTAGCGCGGGGTTCACCCCATCCGAGGCAAGCTCTCTCGG 6960
 DB |||||
 QY 6961 GAGCCAGCTATCGGACCATCGCTGCGAGCCACTGCAACCCAGCGGAGAGAGAGAGAGAG 7020
 DB |||||
 QY 6961 GAGCCAGCTATCGGACCATCGCTGCGAGCCACTGCAACCCAGCGGAGAGAGAGAGAGAG 7020
 DB |||||
 QY 7021 TGTGGACATGCTGATGCTAACTGTTTCATGGGGGCGATGCTGCTCGATAGAGTCTGG 7080
 DB |||||
 QY 7021 TGTGGACATGCTGATGCTAACTGTTTCATGGGGGCGATGCTGCTCGATAGAGTCTGG 7080
 DB |||||
 QY 7081 GTCCAAAGTGGTCTGTTCTGAGCTCTCTCGAACCAATGCTCGAAGAGAGAGAGAGAGAG 7140
 DB |||||
 QY 7081 GTCCAAAGTGGTCTGTTCTGAGCTCTCTCGAACCAATGCTCGAAGAGAGAGAGAGAGAG 7140
 DB |||||
 QY 7141 GCCTTCGATACCATCAGAAATACATGCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAG 7200
 DB |||||
 QY 7141 GCCTTCGATACCATCAGAAATACATGCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAG 7200
 DB |||||
 QY 7201 CTGGGACAGGCTGATTAACAACCCAGCTGTTGGAATCGTGGAAAGAGAGAGAGAGAGAGAG 7260
 DB |||||
 QY 7201 CTGGGACAGGCTGATTAACAACCCAGCTGTTGGAATCGTGGAAAGAGAGAGAGAGAGAGAG 7260
 DB |||||
 QY 7261 ACCGGCCACTGTTGGCGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7320
 DB |||||
 QY 7261 ACCGGCCACTGTTGGCGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7320
 DB |||||
 QY 7321 AAGGAGCGCGGACAGTGGGCTTAAGTAGGAGTCCATAGGAGATGCTCTTCAACAGCT 7380
 DB |||||
 QY 7321 AAGGAGCGCGGACAGTGGGCTTAAGTAGGAGTCCATAGGAGATGCTCTTCAACAGCT 7380
 DB |||||
 QY 7381 GGCCATTAAGTCTTGGCGAGCGGCTTCCAGGCGGATTCAGGCTTTCACGCGGCG 7440
 DB |||||
 QY 7381 GGCCATTAAGTCTTGGCGAGCGGCTTCCAGGCGGATTCAGGCTTTCACGCGGCG 7440
 DB |||||
 QY 7441 GGCGCTGCGGATTCGCGGAGTCCAGAGCTCTCTGATGAGTGGCTTTCGCGAGAGAG 7500
 DB |||||
 QY 7441 GGCGCTGCGGATTCGCGGAGTCCAGAGCTCTCTGATGAGTGGCTTTCGCGAGAGAG 7500
 DB |||||
 QY 7501 TTCCATCTCTTCCATGCGGCTTCCAGGCGGAGCTTGGAGATCCAGAGCTGAGAGCTGAG 7560
 DB |||||
 QY 7501 TTCCATCTCTTCCATGCGGCTTCCAGGCGGAGCTTGGAGATCCAGAGCTGAGAGCTGAG 7560
 DB |||||
 QY 7561 GAGGTAGAGCGGCGGCTTCCAGGCGGAGCTTCCAGGCGGAGCTTCCAGGCGGAGCTGAG 7620
 DB |||||
 QY 7561 GAGGTAGAGCGGCGGCTTCCAGGCGGAGCTTCCAGGCGGAGCTTCCAGGCGGAGCTGAG 7620
 DB |||||
 QY 7621 GTCTGCTGCTTCTGCTTCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7680
 DB |||||
 QY 7621 GTCTGCTGCTTCTGCTTCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7680
 DB |||||
 QY 7681 CTGGAGCGGCGCTTAAATTAATCTCTTGTAGTCCGAGAGAGAGAGAGAGAGAGAGAGAGAG 7740
 DB |||||
 QY 7681 CTGGAGCGGCGCTTAAATTAATCTCTTGTAGTCCGAGAGAGAGAGAGAGAGAGAGAGAGAG 7740
 DB |||||
 QY 7741 CTTGAGCAACTCTCTGTTGGGATATCAACAAAGAGTGTACTGTACCAACAAAGAGAGAGAG 7800
 DB |||||
 QY 7741 CTTGAGCAACTCTCTGTTGGGATATCAACAAAGAGTGTACTGTACCAACAAAGAGAGAGAG 7800
 DB |||||
 QY 7801 CTCACTAAGGCTTAAAGAGTAACTTTTGTAGTACGAGTGTCTGCTGCTCTCTCTCTCTCTCT 7860
 DB |||||
 QY 7801 CTCACTAAGGCTTAAAGAGTAACTTTTGTAGTACGAGTGTCTGCTGCTCTCTCTCTCTCTCT 7860
 DB |||||
 QY 7861 CTGAGTCTTAAAGGACATTAAGCTAGCGGCTTCCAGGCTCAGGAGAGAGAGAGAGAGAGAG 7920
 DB |||||

Db	3481		AGTGGTGAATGACCGGGGCTGACAGACAGAAACAGCGCGGGAAGTCCAAATCTCTGTC	3540
Qy	3541		CACAGTCACCTCAGTCTTCCTCGGAACATCCATCTCGGGGTTTGTGGACTCTCTACCA	3600
Db	3541		CACAGTCTCAGTCTTCCTCGGAACATCTCGGGGTTTGTGGACTCTCTACCA	3600
Qy	3601		TGGAGCTGGCAACAAGACTCTGGCGGCTCACCGGGTCCGGTCAACAGATGACTCCAG	3660
Db	3601		CGGAGCTGGCAACAAGACTCTAGCGGCTTACGGGTCGGGTCAACAGATGACTCCAG	3660
Qy	3661		TGCTGAGGGGACTTAGTAGGTGGCCAGCCCTCGGGTCCGGTCAACAGATGACTCCAG	3720
Db	3661		TGCTGAGGGGACTTAGTAGGTGGCCAGCCCTCGGGTCCGGTCAACAGATGACTCCAG	3720
Qy	3721		CAGCTGGAGCGTCCAGCTGACCTGCTGACCGGTCAGCGGAACGCTGATCCCGGCTCG	3780
Db	3721		CAAGTGGAGCGTCCAGCTGACCTGCTGACCGGTCAGCGGAACGCTGATCCCGGCTCG	3780
Qy	3781		AAGACGGGGACAAACCGGGAGCGCTACTCTCCCGGAGACCTCTTCCACCTTGAAGGG	3840
Db	3781		GAGACGGGGACAAACCGGGAGCGATGCTCTCCCGGAGACCTCTTCCACCTTGAAGGG	3840
Qy	3841		GTCTCAGGAGCGCGTGTATGCCCGAGGGCCAGCGTGTCCGAGTCTTCCGGGCGC	3900
Db	3841		GTCTCAGGAGCGCGTGTATGCCCGAGGGCCAGCGTGTCCGAGTCTTCCGGGCGC	3900
Qy	3901		TGCTGCTCTCGGGGCTGCTAGTCCATAGATTTATCCCGTGGAGACACTGCACAT	3960
Db	3901		TGCTGCTCTCGGGGCTGCTAGTCCATAGATTTATCCCGTGGAGACACTGCACAT	3960
Qy	3961		CGTCACCGCTGCCACCTTTAGTGACACAGCACACACCTGCTGTGCGCCCGACCTTA	4020
Db	3961		TGTTAAGAGTCTCCACTTTAGTGACACAGCACACACCTGCTGTGCGCCCGACCTTA	4020
Qy	4021		TCAGTCCGGTACTGTCATCCCGGCTGCGAGTGGAAAGAGCACCAAGTCTCTGTGCG	4080
Db	4021		TCAGTCCGGTACTGTCATCCCGGCTGCGAGTGGAAAGAGCACCAAGTCTCTGTGCG	4080
Qy	4081		ATATGCTGCTCGGGGTATAGTGCTAGTGTCTTAATCCCTCAGTGGCTGCCACCTGGG	4140
Db	4081		GTATGCTGCTCGGGGTATAGTGCTAGTGTCTTAATCCCTCAGTGGCTGCCACCTGGG	4140
Qy	4141		GTTTGGGGCTACTGTCTAAGGCACATGGCATCAATCCCAACATTTAGGACTGGAGTCAG	4200
Db	4141		GTTTGGGGCTACTGTCTAAGGCACATGGCATCAATCCCAACATTTAGGACTGGAGTCAG	4200
Qy	4201		GACTGTGACACCGGGGCGCCATCAGTACTCCATATGGCAAAATCTCCCGGATGG	4260
Db	4201		GACCGTATGACCGGGGAGCCATCAGTACTCCATATGGCAAAATTTCTCGCCGATGG	4260
Qy	4261		GGGCTGTGGGGCGGCGCTACGACATCATATGTGATGATGCGCTGCGTGAATGC	4320
Db	4261		GGGCTGTGGGGCGGCGCTACGACATCATATGTGATGATGCGCTGCGTGAATGC	4320
Qy	4321		TACACCATCTTGGGATCGGAACAGTCTTGTATCAGAGAGACAGCTGGGTCAGACT	4380
Db	4321		TACCTCATCTTGGGATCGGAACAGTCTTGTATCAGAGAGACAGCTGGGTCAGACT	4380
Qy	4381		AACCTGTGCTGCTACAGCTACGCGGCTGAGTCAAGTCAAGAGAGACAGCTGGGTCAG	4440
Db	4381		AACCTGTGCTGCTACAGCTACGCGGCTGAGTCAAGTCAAGAGAGACAGCTGGGTCAG	4440
Qy	4441		GGAGGTGGCGCTTGGGAGAGGGGAGATCCCGTCTTATGGAGGGGAGTCCCGTGTG	4500
Db	4441		AGAGGTAGGCTTGGGAGAGGGGAGATCCCGTCTTATGGAGGGGAGTCCCGTGTG	4500
Qy	4501		TTACATCAAGGAGAGAGATCTGATCTCTGCAATCAAGAAAGTGTACAGGCT	4560
Db	4501		CTGCATCAAGGAGAGAGACCTGATTTCTGCAATCAAGAAAGTGTACAGGCT	4560
Qy	4561		CGCGGCGGCTTGGGGTATGGGCTTGAATCAGTGGGCTACTACAGAGGGTGGACGT	4620
Db	4561		CGCGGCGGCTTGGGGTATGGGCTTGAATCAGTGGGCTACTACAGAGGGTGGACGT	4620
Db	4561		CGCGGCGGCTTGGGGGATCGGCTTGAATCCGCTGCGATACTATAGAGGGTGGACGT	4620
Qy	4621		CTCCGTAATACCAACTCAGGAGAGAGTGTGTGCTGCGCACCGAGCGCTCATGACAGG	4680
Db	4621		CTCCATATACCAAGCTCAGGAGAGTGTGTGCTGCGCACCGAGCGCTCATGACAGG	4680
Qy	4681		GTATCTGGGAGCTTTGACTCTCGTGTATCGACTGCAACGCTAGCGGTCACTCAAGTGT	4740
Db	4681		GTATCTGGGAGCTTTGACTCTCGTGTATCGACTGCAACGCTAGCGGTCACTCAAGTGT	4740
Qy	4741		CTTCAGTTTAGACCCCAATTTACCNATACCAACAGATGCTCCCTCAAGAGCGCTGTCT	4800
Db	4741		CTTCAGCTTGGAGCCCGGCTTCACTATACCAACAGACTGTCCCAACAGAGCGCTGTCT	4800
Qy	4801		AGTAGCCAGCGCGGGGTCGACAGGCTAGGAGAGACTGGGCTTTATAGGTATGTTTC	4860
Db	4801		AGTAGCTAGCGCGCGGGGTCGACAGGCTAGGAGAGAGAGGAGCTTATAGGTATGTTTC	4860
Qy	4861		CACCTGTGAGCGAGCGCTCAGGAATGTTTGCAGTGTGTGTGCTCTGTGAGTGTCTACG	4920
Db	4861		CACCTGTGAGCGAGCGCTCAGGAATGTTTGCAGTGTGTGTGCTCTGTGAGTGTCTACG	4920
Qy	4921		AGGGGCGCATGCTATGAGTCTCACCATCGGAGACCAAGTGTGAGGCTCAGGCGTATTT	4980
Db	4921		AGGGGCTGCTGTGATCGATCTCACACAGCGGAGACCAAGTGTGAGGCTCAGGCGTATTT	4980
Qy	4981		CACACGCGCGGCTTCCCTGCTGTCGACAGACCATCTTGAGTGTGGGAGCGATTTTCAC	5040
Db	4981		CACACGCGCGGCTTCCCTGCTGTCGACAGACCATCTTGAGTGTGGGAGCGATTTTCAC	5040
Qy	5041		CGGCTCTCACACATAGATGCCACTTCTCTTCCCAACAAAGCAATCCGGGGGAAATTT	5100
Db	5041		CGGCTCTCACACATAGATGCCACTTCTCTTCCCAACAAAGCAATCCGGGGGAAATTT	5100
Qy	5101		CGGATATCTTAAACAGCTTACAGGCTACAGTGTGCGTACAGGCGCAAGCGCCCCCGGTC	5160
Db	5101		CGGATATCTTAAACAGCTTACAGGCTACAGTGTGCGTACAGGCGCAAGCGCCCCCGGTC	5160
Qy	5161		CTGGAGCGCTCATGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	5220
Db	5161		CTGGAGCGCTCATGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	5220
Qy	5221		TCTCTGTACCGCTTGGGCTCTGTGTACCAACAGAGTCAAGCTTCCCTGAGTGTGCTG	5280
Db	5221		TCTCTGTACCGCTTGGGCTCTGTGTACCAACAGAGTCAAGCTTCCCTGAGTGTGCTG	5280
Qy	5281		ATACATCGGCACTGTGATGCAAGCGACCTTGTGAGTGTGATGAGTGTGCTGCTGCTG	5340
Db	5281		ATACATCGGCACTGTGATGCAAGCGACCTTGTGAGTGTGATGAGTGTGCTGCTGCTG	5340
Qy	5341		AGGGGAGCTTGGGCGGCTGCGCGCTTGTGCTGCGGAGCGGGTGTGCTGCTGCTGCTG	5400
Db	5341		AGGGGAGCTTGGGCGGCTGCGCGCTTGTGCTGCGGAGCGGGTGTGCTGCTGCTGCTG	5400
Qy	5401		CGGCGCTTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG	5460
Db	5401		CGGCGCTTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG	5460
Qy	5461		GGCTTTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG	5520
Db	5461		GGCTTTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG	5520
Qy	5521		GATAGCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG	5580
Db	5521		GATAGCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG	5580
Qy	5581		TCAAGACATACCAACCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG	5640
Db	5581		TCAAGACATACCAACCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG	5640
Qy	5641		ACACATGTGGAATCTTCAATGTGAGGCTTCAATGTGAGGCTTCAATGTGAGGCTT	5700
Db	5641		ACACATGTGGAATCTTCAATGTGAGGCTTCAATGTGAGGCTTCAATGTGAGGCTT	5700

1321 GCGGCGGCTACCATGATCTTGGCGGTACGGATGCGTGTGCGGAGGTCAATATGACAT 1380
 1321 GCGGCGGCGACCATGATCTTGGCGGTACGGATGCGGTCGCGGTCCCGAGGTCAATATGACAT 1380
 1381 CATTTAGCGGGCTCATTTGGGGCGTCAATGTTGGGCTTTGGGCTACTTCTCTATGAGGAGC 1440
 1381 CGTTAGCGGGCTCATTTGGGGCGTCAATGTTGGGCTTTGGGCTACTTCTCTATGAGGAGC 1440
 1441 GTGGGCGAAAGTCTGTTGATCTTCTGTTGGGCGGCGGGGTGAGCGGCGACCATATAC 1500
 1441 GTGGGCGAAAGTCAATTTGATCTTCTGTTGGGCGGCGGGGTGAGCGGCGACCATATAC 1500
 1501 TGTGGGGGTCTGCGCGGAGACACCGGGCGCTTCAACAGCTTATTTGACATGCGGCC 1560
 1501 CGTTGAGGGCGCTTGTGACGCTTCAACCAAGCTGTTGCGGGCTGTTGAGCCATGGGCC 1560
 1561 CAGGCGAAATCCAGCTCGTTAAACCAATGCGAGCTGGCAATCAACCGCACCGCCCT 1620
 1561 TCAGCAGAACATTCAGCTCAATTAACCAACCAAGCGAGTTGGCAATCAACCGTACTGCTT 1620
 1621 GAATGCAATGACTCTTGGCAGACCGGCTTTATGCGGTCTCTGTTCTACACCCACAGCTT 1680
 1621 GAATGCAATGACTCTTGGCAGACCGGCTTTCTGCGGGCTTGTCTTACACCAACCGCTT 1680
 1681 CAATCGTCAAGGATGTCGGAACGCAATGTCGCGCTGCGGAGTATCGAGGCTTTCGGGT 1740
 1681 TAACTGTCAGGGTGTTCAGGGCGCTGTCGCGCTGCGGCAATCGAGGCTTTCGGAT 1740
 1741 GGGATGGGGCGCTTGAATATGAGGATATGCAATCAATCCAGAGATATGAGACCTTA 1800
 1741 AGGGTGGGCGACCTTACAGTACGAGGATATGTCACCAATCCAGAGGATATGAGGCGTA 1800
 1801 TTGCTGCACTACCCAGGAGGCTGTCGCGGTGCTTCCGCAAGACTGTTGTCGCC 1860
 1801 CTGCTGGCACTTACCCCGGAGGCTGTCGCGGTGCTTCCGCAAGACTGTTGTCGCC 1860
 1861 AGTGTACTGTTTCAACCCAGGAGGCTGTCGCGGTGCTTCCGCAAGACTGTCGAGCGCC 1920
 1861 AGTGTACTGTTTCAACCCAGGAGGCTGTCGCGGTGCTTCCGCAAGACTGTCGAGCGCC 1920
 1921 CACTTACAGTGGGGGAGATGAGCAGATGCTTCTTCAATGAACAGCACTCGACAC 1980
 1921 CACTTACAGTGGGGGAGATGAGCAGATGCTTCTTCAATGAACAGCACTCGACAC 1980
 1981 GCTGGGTCAATGTTGCGCTGCACTGAGTGAATCTTCTGCGCTACACCAAGACTTGGG 2040
 1981 GCAAGGCTCAATGTTGCGCTGCACTGAGTGAATCTTCTGCGCTACACCAAGACTTGGG 2040
 2041 CGCACCACTTCCGCTACTAGAGTGAATCTTCAACCGGAGGAGGCTGTTGTCGCCAC 2100
 2041 CGGCCACCTTCCGCTACTAGAGTGAATCTTCAACCGGAGGAGGCTGTTGTCGCCAC 2100
 2101 GGAATGTTTGGGAGGATCTGATGCACTTATTAAGTGTGTTCTGGGCGCTGGCT 2160
 2101 GGAATGTTTGGGAGGATCTGATGCACTTATTAAGTGTGTTCTGGGCGCTGGCT 2160
 2161 CAGCCAGGCTGCTGATGCACTTACAGGCTTGGCACTTACCCCTGCAAGTAA 2220
 2161 CAGCCAGGCTGCTGATGCACTTACAGGCTTGGCACTTACCCCTGCAAGTAA 2220
 2221 CTATACCATCTTCAAAATAGAGTGTATGTTGGAGGAGGCTTGGAGCAGAGGCTCAGCGTGC 2280
 2221 TTTTACCATCTTCAAGATAGAGTGTATGTTGGGAGGCTTGGAGCAGAGGCTCAGCGTGC 2280
 2281 ATGCAATTTCACTGCTGGGATCTGCACTTGGAGGAGCAGAGCAGAGAGTCAATGTC 2340
 2281 ATGCAATTTCACTGCTGGGATCTGCACTTGGAGGAGCAGAGCAGAGAGTCAATGTC 2340
 2341 TCCTTTGTTGCACTTCCAGGAGTGGGCTTTTACCTTGTCTTACTCGGACCTGCC 2400
 2341 TCCTTCTGTGCACTTCCAGGAGTGGGCTTCTGCTCCCTGCACTTCAAGACTTACC 2400

2401 GCGCTTGTGAGTGTCTTCCACTCCACCAAAACATCGTGGAGCTCAATTCATGTA 2460
 2401 GCGCTTGTGCACTTGTCTTCTCCACTCCACCAAAACATCGTGGAGCTCAATTCATGTA 2460
 2461 TGGCCTATCACTGCGCTCAGAAAATACATCGTCCGATGGGAGTGGGTAACTCTTAT 2520
 2461 TGGCCTCTCACTGCTTATCAAAAATACATCGTCCGATGGGAGTGGGTAACTCTTAT 2520
 2521 CTGCTCTTGTAGGAGCGCGGAGGCTTGTGCGCTTGTATGATGCTCATCTTGTGGGCCA 2580
 2521 CTGCTCTTGTAGGAGCGCGGAGGCTTGTGCGCTTGTATGATGCTCATCTTGTGGGCCA 2580
 2581 GSCCGAAGCAGCACTAGAGAGCTGTCTTGTGACGCTGCGAGCGAGCTAGCTGCA 2640
 2581 GSCCGAAGCAGCACTAGAGAGGTTGTGCTGTCTTGCAGCTGCGAGTGGCGCTAACTGCCA 2640
 2641 TGGCTTCTTATTTTGTCTCTTTTGTGCTGCTTGTGCTGCTGCTACATCAAGGCTCGGTACT 2700
 2641 TGGCTTCTTATTTTGTCTCTTTTGTGCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCT 2700
 2701 CCGCTTGTGCTTACTTCTTCTTCTGCTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCT 2760
 2701 CCGCTTGTGCTTACTTCTTCTTCTGCTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCT 2760
 2761 GSCCGAAGCAGGCTTATGCTTATGAGCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 2820
 2761 GSCCGAAGCAGGCTTATGCTTATGAGCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 2820
 2821 GGTATGATGCTTCTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880
 2821 GATATGATGCTTCTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880
 2881 GTGCTGCTGCTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2940
 2881 GTGCTGCTGCTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2940
 2941 TATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3000
 2941 CATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3000
 3001 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3060
 3001 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3060
 3061 AGTGTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3120
 3061 GCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3120
 3121 CATGCGAAGGCTTCTGCGGGGGGAGGTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3180
 3121 TTTGCTGAGGAGCTGCTGCGGGGGGAGGTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3180
 3181 GTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3240
 3181 GTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3240
 3241 CCGTGGGAGCCTGCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300
 3241 CCGTGGGAGCCTGCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300
 3301 CATTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3360
 3301 CATCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3360
 3361 GCGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420
 3361 GCGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420
 3421 GAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3480
 3421 GAAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3480
 3481 AGTGTGTGAGCATGAGCGGGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3540

181 AAGACTGGGTCTTTCTTGGATAAACCCACTCTATGCCCGGCCATTGGGCGTGCCCCG 240

Copied from 09540843 on 05/19/2004

7261 ACCGGCCTGTTGGCGGCTGCGCTCTCCCGCCCTTAAGAAAACCCGACGCTCCCGC 7320
 7321 AAGAGCGCCGAGCAGTGGCCCTAAGTAGAGACTCCATAGGAGATGCCCTTCAACAGCT 7380
 7321 AAGAGCGCCGAGCAGTGGGCTCTGAGTAGAGACTCCATAGCAGATGCCCTCAACAGCT 7380
 7381 GGCATTAAAGTCTTTGGCAGAGCCCGCCCAAGCGGGAATTCAGGCTTTTCCACGGGGC 7440
 7381 GGCATCAAGTCTTTGGCAGAGCCCGCCCAAGCGGGAATTCAGGCTTTTCCACGGGGC 7440
 7441 GGGCGCTGCGGATTCGCGAGTCAAGCGCTCTGATGAGTGGCGCTTTTCGAGACAG 7500
 7441 GAGCGAGCGGATTCGCGAGTCAAGCGCTCTGATGAGTGGCGCTTTTCGAGACAG 7500
 7501 TTCATCTCTTCATGCGCCCTTCGAGGGGAGCTTGGAGATCCAGACCTGAGCTGA 7560
 7501 TTCATCTCTTCATGCGCCCTTCGAGGGGAGCTTGGAGATCCAGACTTGGAGCTGA 7560
 7561 GCAGGTAGAGCCCAACCCCGCCCAAGGGGGGGTGGCAGCTCCCGGCTCGGACTGGG 7620
 7561 GCAGGTAGAGCTTCAACCTCCCGCCCAAGGGGGGGTGGTAAACCCCGGCTCAGGCTCGG 7620
 7621 GTCGTGCTACTTCTCCGAGAGAGAGTCTCGTGTGCTGCTGCTCCATGCTACTC 7680
 7621 GTCGTGCTACTTCTCCGAGAGAGAGTCTCGTGTGCTGCTGCTCCATGCTACTC 7680
 7681 CTGACCGGGCTCTAATACTCTTCTAGTCCCGAGAGAGAGTTCACGATTAACCC 7740
 7681 CTGACCGGGCTCTAATACTCTTCTAGTCCCGAGAGAGAGTTCACGATTAACCC 7740
 7741 CTTGAGCAACTCCCTGTTGGGATATCAACAAGGTTGCTGTACCAACAAGAGCGC 7800
 7741 CTTGAGCAACTCCCTGTTGGGATATCAACAAGGTTGCTGTACCAACAAGAGCGC 7800
 7801 CTCATAAGGGCTAAAGGTAACTTTTGTAGTAGATGCAAGTGTGCTGACTCTACTAG 7860
 7801 CTCATAAGGGCTAAAGGTAACTTTTGTAGTAGATGCAAGTGTGCTGACTCTACTAG 7860
 7861 CTCAGTCTTAAGGACATTAAGCTAGCGGCTCCCAAGTCAACGAGGCTCTCAACAT 7920
 7861 CTCAGTCTTAAGGACATTAAGCTAGCGGCTCCCAAGTCAACGAGGCTCTCAACAT 7920
 7921 GGAGGGCTTGGCAGTTAAACCCCAACCTTCTGCAAGATCTAATATATGGGTTGGGGC 7980
 7921 AGAGGGCTTGGCAGTTAACTCCACCCCTCTGCAAGATCCCAAGTATGGGTTGGGGC 7980
 7981 TAAGGAGTTCGAGCTTGTTCGGGAGGCGGTTAACCATCAAGTCCGTGTGAAGGA 8040
 7981 TAAGGAGTTCGAGCTTGTTCGGGAGGCGGTTAACCATCAAGTCCGTGTGAAGGA 8040
 8041 CCTCTGGAGCTCAGAAACCAATTTCCCAACCAATTTATGSCCAAAATGAGGTGT 8100
 8041 CCTCTGGAGCTCAGAAACCAATTTCCCAACCAATTTATGSCCAAAATGAGGTGT 8100
 8101 CTGCGTGGACCCCAAGGGGGCAAGAGCAGCTCGCTTATCGTTTACCCCTGACCT 8160
 8101 CTGCGTGGACCCCAAGGGGGCAAGAGCAGCTCGCTTATCGTTTACCCCTGACCT 8160
 8161 GCGGCTCAGGGTCTGCGAGAGATGGCCCTTTATGACATTAACAACCACTTCTCAGGC 8220
 8161 GCGGCTCAGGGTCTGCGAGAGATGGCCCTTTATGACATTAACAACCACTTCTCAGGC 8220
 8221 GGTATGGGGCTCTTATGATTCAGATTTCCCGCTCAGCGGTAGAGTTTCTCT 8280
 8221 GGTATGGGGCTCTTATGATTCAGATTTCCCGCTCAGCGGTAGAGTTTCTCT 8280
 8281 GAAAGCATGGCGGAAAGAGACCTTATGGGTTTTTGTATGATACCCGATGCTTTGA 8340
 8281 GAAAGCATGGCGGAAAGAGACCTTATGGGTTTTTGTATGATACCCGATGCTTTGA 8340
 8341 CTCACCGTCTCAGAGAGACATCAGAGCTGAGAGTCCATATATCGGGCTCTCCTT 8400
 8341 CTCACCGTCTCAGAGAGACATCAGAGCTGAGAGTCCATATATCGGGCTCTCCTT 8400

8401 GCCGAGGAGGCCACACTGCGCATACACTGCTAACTGAGAGACTTTTACGTGGAGGGCC 8460
 8401 GCCGAGGAGGCCACACTGCGCATACACTGCTAACTGAGAGACTTTTACGTGGAGGGCC 8460
 8461 TATGTTCAACAGCAAGGGCCAACTTGGGGTACAGCGCTTTCGCGGCCAGCGGGGTGCT 8520
 8461 CATGTTCAACAGCAAGGGCCAGACTTGGCGGTACAGCGCTTTCGCGGCCAGCGGGGTGCT 8520
 8521 CACCACTAGCATGGGGAACACCATACATGCTACGTGAAAGCTTACGCGCTTGTAAAGC 8580
 8521 TACCACTAGCATGGGGAACACCATACATGCTATGTGAAGCTTACGCGCTTGTAAAGC 8580
 8581 TGCAAGGATAAATCGCGCCCACAATGCTGTATGCGCGGATGACTTGTGTTGTCATCTCAG 8640
 8581 TGCAAGGATAAATGCGCCCACAATGCTGTATGCGCGGATGACTTGTGTTGTCATCTCAG 8640
 8641 AAGCAAGGGGACCGAGGAGGACGAGCGGAACCTTGAGAGCTTTCACGAGGCTATGACCA 8700
 8641 GAGCAAGGGGACCGAGGAGGACGAGCGGAACCTTGAGAGCTTTCACGAGGCTATGACCA 8700
 8701 GTATTCTGCGCCCTCTCTGCTGAGCCCGCCAGAGCGGAGTATGATCTGAGCTGATAACATC 8760
 8701 GTATTCTGCGCCCTCTCTGCTGAGCCCGCCAGAGCGGAGTATGATCTGAGCTGATAACATC 8760
 8761 TTGCTCTCAAAATGCTGCTGCTGCGGCTGCGGCCCAAGGCGCGCAGATACCTACCTGAC 8820
 8761 TTGCTCTCAAAATGCTGCTGCTGCGGCTGCGGCCCAAGGCGCGCAGATACCTACCTGAC 8820
 8821 CAGAGACCTTACCTCAATCGCCGCGCTGCTGCGGAAACAGATTAGACACTTCCCTGT 8880
 8821 CAGAGACCTTACCTCAATCGCCGCGCTGCTGCGGAAACAGATTAGACACTTCCCTGT 8880
 8881 CAATTATCGCTGGGAAACATCATCCAGTACGCGCCCAAGATCTCTGAGACAGAACTCACTT 8940
 8881 CAATTATCGCTGGGAAACATCATCCAGTACGCGCCCAAGATCTCTGAGACAGAACTCACTT 8940
 8941 GATGACACACTTCTCTCCATCTCATGCTCAAGACACGCTGAGACAGAACTCACTT 9000
 8941 GATGACACACTTCTCTCCATCTCATGCTCAAGACACGCTGAGACAGAACTCACTT 9000
 9001 TGAGATGATCGAGCGGTGTATCTCCGTGAGTCCCTTGGAGCTCCCGAGCTATATTAAG 9060
 9001 TGAGATGATCGAGCGGTGTATCTCCGTGAGTCCCTTGGAGCTCCCGAGCTATATTAAG 9060
 9061 GTTACATGCTGAGCGCTTTTCTCTGACACATACACTTCCCAAGATCTGAGACAGCGGT 9120
 9061 GTTACATGCTGAGCGCTTTTCTCTGACACATACACTTCCCAAGATCTGAGACAGCGGT 9120
 9121 GGCTTCAGCGCTCAGAAAACTTGGGCGCCACCCCTCAGAGCGTGAAGAGCCGGGCA 9180
 9121 GGCTTCAGCGCTCAGAAAACTTGGGCGCCACCCCTCAGAGCGTGAAGAGCCGGGCA 9180
 9181 TGAGTTCAGCGCGCTCCTCATCTCCGTGGGGGAGAGCGCGCTTTCGCGCTGATATCT 9240
 9181 TGAGTTCAGCGCGCTCCTCATCTCCGTGGGGGAGAGCGCGCTTTCGCGCTGATATCT 9240
 9241 CTTCAATGGCGGTCAAGACCAAGCTCAAACTCACTTCCATTCGCGGAGCGCGCTTCT 9300
 9241 CTTCAATGGCGGTCAAGACCAAGCTCAAACTCACTTCCATTCGCGGAGCGCGCTTCT 9300
 9301 GGATTTATCAGCTGGTTTACCGCTCGCGCGCGCGGGGCGACATTTATCAGAGCTGTC 9360
 9301 GGATTTATCAGCTGGTTTACCGCTCGCGCGCGCGGGGCGACATTTATCAGAGCTGTC 9360
 9361 GCGTGCAGCGCGCTTATGCTCTTTGCGCTTACTCTACTTTTGTAGGGGTAGGCT 9420
 9361 GCGTGCAGCGCGCTTATGCTCTTTGCGCTTACTCTACTTTTGTAGGGGTAGGCT 9420
 9421 TTTTCTACTCTCCCGCTGGGTAGAGCGGACACATTAAGTACCTCCATACCTACTGTC 9480
 9421 TTTTCTACTCTCCCGCTGGGTAGAGCGGACACATTAAGTACCTCCATACCTACTGTC 9480

5101 CGCATCTTAAACAGCCTACAGGCTACAGTGTGCGCTAGGGCCAAAGCCCCCCCCCGTC 5160
 5101 CGCATCTTAAACAGCCTACAGGCTACAGTGTGCGCTAGGGCCAAAGCCCCCCCCCGTC 5160
 5161 CTGGGACGTCATGTGGAAGTGTGTGATCTGATCTCAAGCCCAACATCTCGTGGGCCCCACACC 5220
 5161 CTGGGACGTCATGTGGAAGTGTGTGATCTGATCTCAAGCCCAACATCTCGTGGGCCCCACACC 5220
 5221 TCTCTGTACCGCTTGGGCTCTGTTACCAACAGGCTCACCTCACACATCTCGTGAAGAA 5280
 5221 TCTCTGTACCGCTTGGGCTCTGTTACCAACAGGCTCACCTCACACATCTCGTGAAGAA 5280
 5281 ATACATCGCACCTGATGCAAGCCGACCTTGAGGTCATGACAGACATGAGTCTTTGGC 5340
 5281 ATACATCGCACCTGATGCAAGCCGACCTTGAGGTCATGACAGACATGAGTCTTTAGC 5340
 5341 AGGGGAGTCTTGGGCGCTGCGCGGTATTCCTTGGGACCGCGGTGTGTTGATCAT 5400
 5341 TGGGGGGGCTCTTGGCGCGCTGCGCGGTATTCCTTGGGACCGCGGTGTGTTGATCAT 5400
 5401 CGGCGCTTGCATTAACAGGAGTGTGCTTGGGCGGTCTCATTTGAAGAGGGGCGAGC 5460
 5401 CGGCGCTTGCATTAACAGGAGTGTGCTTGGGCGGTCTCATTTGAAGAGGGGCGAGC 5460
 5461 GGCTTTTGTAGATGGAGGATGTGCTTGGGCGGTCTCATTTGAAGAGGGGCGAGC 5520
 5461 GGCTTTTGTAGATGGAGGATGTGCTTGGGCGGTCTCATTTGAAGAGGGGCGAGC 5520
 5521 GATAGCGGAGTCTGAAGTCCAAGATCCAAGGCTTATTGACAGAGGCTTCCAACAGC 5580
 5521 GATAGCGGAGTCTGAAGTCCAAGATCCAAGGCTTATTGACAGAGGCTTCCAACAGC 5580
 5581 TCAAGACATACACCTGTCAGGCTTCATGCGGCTTGGGCGGCTGAGTGAACATCTTGGGCGAA 5640
 5581 TCAAGACATACACCTGTCAGGCTTCATGCGGCTTGGGCGGCTGAGTGAACATCTTGGGCGAA 5640
 5641 ACATGTTGAACTTCATTTAGCGGCTTCAATACCTCGAGGAGTCAACACTGCGAGC 5700
 5641 ACATGTTGAACTTCATTTAGCGGCTTCAATACCTCGAGGAGTCAACACTGCGAGC 5700
 5701 GAACCTGCGAGTTCATGATGCGGCTTCAAGTGGGCGCTTCAACAGTCCGTTGTCAC 5760
 5701 GAACCTGCGTGTAGTTCATGATGCGGCTTCAAGTGGGCGCTTCAACAGTCCGTTGTCAC 5760
 5761 AAGCACATCTCTTCAAGATTTTGGGGGCTGCTAGCATCTCAATTTGCCACAC 5820
 5761 TAGCACATCTCTTCAAGATTTTGGGGGCTGCTAGCATCTCAATTTGCCACATCTCT 5820
 5821 CGCGGGGCGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGG 5880
 5821 CGCGGGGCGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGG 5880
 5881 CTTAGGTAAGTGTAGTGGACATCTTGGGAGGATGTTGGGAGGATTTGGGGGCTCT 5940
 5881 CTTGGGTAAAGTGTCTGTTGACATCTTGGGAGGATGTTGGGAGGATTTGGGGGCTCT 5940
 5941 CGTGGCATCAAGATCTGTTGGGAGGAGGCTTGGGAGGATGTTGGGAGGATTTGGGAGG 6000
 5941 CGTGGCATCAAGATCTGTTGGGAGGAGGCTTGGGAGGATGTTGGGAGGATTTGGGAGG 6000
 6001 GCTTGAATTTCTGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 6060
 6001 GCTTGAATTTCTGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 6060
 6061 CCGACATGTTGGGAGCGGGGAGGCGCTGTTCAATGATGATGATGATGATGATGATGATG 6120
 6061 CCGACATGTTGGGAGCGGGGAGGCGCTGTTCAATGATGATGATGATGATGATGATGATG 6120
 6121 TTTCCAGGAGGAAATCAAGTCTGCGGCGGCGGCTTGGGCTTGGGCTTGGGCTTGGGCT 6180
 6121 TTTCCAGGAGGAAATCAAGTCTGCGGCGGCGGCTTGGGCTTGGGCTTGGGCTTGGGCT 6180
 6181 TGTGACCAACTACTTGGCTCCCTTACATAACAGGCTGCTCAGAAAGACTCCCAACTG 6240

6181 TGTGACCAACTACTTGGCTCCCTTACATAACAGGCTGCTCAGAGACTCCCAACTG 6240
 6241 GATTAAGGAGTCTGCGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGG 6300
 6241 GATTAAGGAGTCTGCGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGG 6300
 6301 GGTGGACCATCTTAAACAGACTTTAAATAATTTGGGTGACCTCCAAATTTATTTCCCAAGAT 6360
 6301 GGTGGACCATCTTAAACAGACTTTAAATAATTTGGGTGACCTCCAAATTTATTTCCCAAGAT 6360
 6361 GCGCGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGG 6420
 6361 GCGTGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGG 6420
 6421 CATCATGACCAACAGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 6480
 6421 TATCATGACCAACAGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 6480
 6481 CATGAGATCAAGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGG 6540
 6481 CATGAGATCAAGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGG 6540
 6541 TTTTACACGAGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGG 6600
 6541 TTTTACACGAGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGG 6600
 6601 GAGGCTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 6660
 6601 GAGGCTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 6660
 6661 AGGACTTACCATGATTAAGTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 6720
 6661 AGGACTTACCATGATTAAGTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 6720
 6721 CTGGGTGAGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 6780
 6721 CTGGGTGAGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 6780
 6781 TGAGTCTGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 6840
 6781 TGAGTCTGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 6840
 6841 CCTGAAACCGACACAGAGTATGATGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGG 6900
 6841 TCTGAACTTACACAGAGTATGATGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGG 6900
 6901 GGAGACTGAGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 6960
 6901 GGAGACTGAGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 6960
 6961 GAGGAGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGG 7020
 6961 GAGGAGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGG 7020
 7021 TGTGAGCATGTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 7080
 7021 TGTGAGCATGTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 7080
 7081 GTTCCAAAGTGGTCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGG 7140
 7081 GTTCCAAAGTGGTCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGG 7140
 7141 GCTTTCGATACCATGAGATATGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 7200
 7141 GCTTTCGATACCATGAGATATGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 7200
 7201 CTGGGACGCGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 7260
 7201 CTGGGACGCGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 7260
 7261 ACCGCGCACTGTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 7320

Db 2521 CCTGCTTAGCGAGCCAGGTTTCGCCCTCTCTATGGANGCTCATCTTGTGGGCCA 2580
 Qy 2581 GGCCGAAGCAGCTAGAGAGCTGTGTCATCTTGCAGCTGCGAGCGCAGCTAGCTGCA 2640
 Db 2581 GGCCGAAGCAGCTAGAGAGCTGTGTCATCTTGCAGCTGCGAGCGCAGCTAGCTGCA 2640
 Qy 2641 TGGCTTCTCTATATTTTGTCTATCTTTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 2700
 Db 2641 TGGCTTCTCTATATTTTGTCTATCTTTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 2700
 Qy 2701 CCCCCTAGCTACCTATCTTCT 2760
 Db 2701 CCCCCTAGCTACCTATCTTCT 2760
 Qy 2761 GCCCAACAGCTTATGCTTATGAGCATCTGTGCTATGCGCATAGGAGCGCTCTGCT 2820
 Db 2761 GCCCAACAGCTTATGCTTATGAGCATCTGTGCTATGCGCATAGGAGCGCTCTGCT 2820
 Qy 2821 GGTAATGATCACTCTCTTTTACTCTCACCCCGGGTATAAGACCCTTTCTCAGCCGGTTTTT 2880
 Db 2821 GGTAATGATCACTCTCTTTTACTCTCACCCCGGGTATAAGACCCTTTCTCAGCCGGTTTTT 2880
 Qy 2881 GTGGTGTGTGCTATCTTCTGACCTTGGGGGAAGCTATGCTCAGGAGTGGGCAACC 2940
 Db 2881 GTGGTGTGTGCTATCTTCTGACCTTGGGGGAAGCTATGCTCAGGAGTGGGCAACC 2940
 Qy 2941 TATGCAAGTGGGGTGGCGTATGATGATATGCGCATATGCGCGTGCCTATTTCTACCCAG 3000
 Db 2941 TATGCAAGTGGGGTGGCGTATGATGATATGCGCATATGCGCGTGCCTATTTCTACCCAG 3000
 Qy 3001 TGTGGTGTGTGATCAATCAACCAAGTGGCTTCTTGGCGGTGCTTGGCGCTTACCTCTCTAAA 3060
 Db 3001 TGTGGTGTGTGATCAATCAACCAAGTGGCTTCTTGGCGGTGCTTGGCGCTTACCTCTCTAAA 3060
 Qy 3061 AGTGTCTTGAACGCGGTGCGTACTTCTGTCAGGCTCAGCTCTACTAGAGATGTCAC 3120
 Db 3061 AGTGTCTTGAACGCGGTGCGTACTTCTGTCAGGCTCAGCTCTACTAGAGATGTCAC 3120
 Qy 3121 CATGGCAAGCATCTCGCGGGGCGAGTACGTCAGATGGCGCTACTAGCCCTTGGCAG 3180
 Db 3121 CATGGCAAGCATCTCGCGGGGCGAGTACGTCAGATGGCGCTACTAGCCCTTGGCAG 3180
 Qy 3181 GTGGACTGGCACTTACATCTATGACCACTTCAACCTTATGTCGGATTTGGGCTCTAGTGG 3240
 Db 3181 GTGGACTGGCACTTACATCTATGACCACTTCAACCTTATGTCGGATTTGGGCTCTAGTGG 3240
 Qy 3241 CTTGGGGACCTGGCGGTGCGCTTGAAGCTATCATCTTCACTGTCGATGAGAGAAAGT 3300
 Db 3241 CTTGGGGACCTGGCGGTGCGCTTGAAGCTATCATCTTCACTGTCGATGAGAGAAAGT 3300
 Qy 3301 CATTTCTGGGGAGCGGAGACACTGCTTGTGGGGACATTTTACCGGACTTCCCGTCTC 3360
 Db 3301 CATTTCTGGGGAGCGGAGACACTGCTTGTGGGGACATTTTACCGGACTTCCCGTCTC 3360
 Qy 3361 CGCCCGACTTGTGCGGAGTCTCTCTTGGCCAGCTGATGGCTATACCTTCAAGGGGTG 3420
 Db 3361 CGCCCGACTTGTGCGGAGTCTCTCTTGGCCAGCTGATGGCTATACCTTCAAGGGGTG 3420
 Qy 3421 GAGTCTTCTCGCCCCCATCTCTTAAGCCAGCAGACAGCTGGGCTTTTGGGCAACCAT 3480
 Db 3421 GAGTCTTCTCGCCCCCATCTCTTAAGCCAGCAGACAGCTGGGCTTTTGGGCAACCAT 3480
 Qy 3481 AGTGGTGAAGCATGACCGGGGCGGACAGACAGACAGGCTGGGGAATTCAGTCTCTGTC 3540
 Db 3481 AGTGGTGAAGCATGACCGGGGCGGACAGACAGACAGGCTGGGGAATTCAGTCTCTGTC 3540
 Qy 3541 CACAGTCACTCAGTCTCTTCTCGGAAACATCTCTCGGGGTCTTGTGGACTGTCTACCA 3600
 Db 3541 CACAGTCACTCAGTCTCTTCTCGGAAACATCTCTCGGGGTCTTGTGGACTGTCTACCA 3600
 Qy 3601 TGGAGCTGGCAACAAGACTCTGCGCGCTCAGCGGGTCCGCTCAGCAGATGATCTCCAG 3660
 Db 3601 TGGAGCTGGCAACAAGACTCTGCGCGCTCAGCGGGTCCGCTCAGCAGATGATCTCCAG 3660

Qy 3661 TGCTCAGGGGAGCTTAGTAGGTGCGCCAGCCCTTGGGACTAAATCTTTGGAGCCGTG 3720
 Db 3661 TGCTCAGGGGAGCTTAGTAGGTGCGCCAGCCCTTGGGACTAAATCTTTGGAGCCGTG 3720
 Qy 3721 CACGTGTGAGCGGTGCGCTGTAATCTGCTGTCAGCGGAAACGCTGATGATCATCCCGCTG 3780
 Db 3721 CACGTGTGAGCGGTGCGCTGTAATCTGCTGTCAGCGGAAACGCTGATGATCATCCCGCTG 3780
 Qy 3781 AAGACGCGGGGACAAACCGGGAGCGCTATCTCTCCCGAGACCTCTTCCACCTTGAAGG 3840
 Db 3781 AAGACGCGGGGACAAACCGGGAGCGCTATCTCTCCCGAGACCTCTTCCACCTTGAAGG 3840
 Qy 3841 GTCTCTCAGAGGCGCGGTGCTATGCTCCCGAGGGGCGACGCTGTCGGAGTCTTCCCGGCGAC 3900
 Db 3841 GTCTCTCAGAGGCGCGGTGCTATGCTCCCGAGGGGCGACGCTGTCGGAGTCTTCCCGGCGAC 3900
 Qy 3901 TGTGTGCTCTCGGGCGGTGCTTAAAGTGCATAGATTTTCAATCCCTGCTGAGACACTGACAT 3960
 Db 3901 TGTGTGCTCTCGGGCGGTGCTTAAAGTGCATAGATTTTCAATCCCTGCTGAGACACTGACAT 3960
 Qy 3961 CGTCAAGCGGTCCCCACCTTTAGTGACACAGACACACCTGCTGTCGCCAGACCTA 4020
 Db 3961 CGTCAAGCGGTCCCCACCTTTAGTGACACAGACACACCTGCTGTCGCCAGACCTA 4020
 Qy 4021 TCAGGTGCGGTACTTGTGCTATGCTGCGACTGCGAGTGGAAAGACACAAAGTCTCTGTCG 4080
 Db 4021 TCAGGTGCGGTACTTGTGCTATGCTGCGACTGCGAGTGGAAAGACACAAAGTCTCTGTCG 4080
 Qy 4081 ATATGCTGCTCAGGGGTATAAAGTGTCTAGTGTCTTAAATCCCTCAGTGGCTGCCACCTGG 4140
 Db 4081 ATATGCTGCTCAGGGGTATAAAGTGTCTAGTGTCTTAAATCCCTCAGTGGCTGCCACCTGG 4140
 Qy 4141 GTTTGGGGGTACTTGTCTAAGGACATGCGATCAATCCCAACATTAGGACTGGAGTCA 4200
 Db 4141 GTTTGGGGGTACTTGTCTAAGGACATGCGATCAATCCCAACATTAGGACTGGAGTCA 4200
 Qy 4201 GACTGTGACGACCGGGGCGCCCTCAGCTACTTCCACATATGCAAAATCTCTCGCGATGG 4260
 Db 4201 GACTGTGACGACCGGGGCGCCCTCAGCTACTTCCACATATGCAAAATCTCTCGCGATGG 4260
 Qy 4261 GGGTGTGCGGGCGCGCTACGACATCATCATATGTGATGAATGCCATGCGGTGAGCTC 4320
 Db 4261 GGGTGTGCGGGCGCGCTACGACATCATCATATGTGATGAATGCCATGCGGTGAGCTC 4320
 Qy 4321 TACCAACCATCTTGGCATCGGAAACAGTCTTGTGATCAAGCAGACAGCTGGGGTCA 4380
 Db 4321 TACCAACCATCTTGGCATCGGAAACAGTCTTGTGATCAAGCAGACAGCTGGGGTCA 4380
 Qy 4381 AACTGTGCTGGCTACAGCTACGCCCTTGGGTGAGTCAACCCCGACCCCAACATAGA 4440
 Db 4381 AACTGTGCTGGCTACAGCTACGCCCTTGGGTGAGTCAACCCCGACCCCAACATAGA 4440
 Qy 4441 GGAGGTGGCCCTTGGGCGAGGGGCGAGATCCCTTCTATGGAGGGCGATTTCCCTGTGTC 4500
 Db 4441 GGAGGTGGCCCTTGGGCGAGGGGCGAGATCCCTTCTATGGAGGGCGATTTCCCTGTGTC 4500
 Qy 4501 TTACATCAAGGGAGGAGACATCTGATCTTGTGCCATTTCAAGAAAAAGTGTGACGCT 4560
 Db 4501 TTACATCAAGGGAGGAGACATCTGATCTTGTGCCATTTCAAGAAAAAGTGTGACGCT 4560
 Qy 4561 CGCGGCGCCCTTGGGGGTATGGGCTTGAACCTCAGTGGCATACTACAGAGGGTGGACCT 4620
 Db 4561 CGCGGCGCCCTTGGGGGTATGGGCTTGAACCTCAGTGGCATACTACAGAGGGTGGACCT 4620
 Qy 4621 CTTCCGTAATACCACTCAGGGGAGAGTGTGCTGCTCGCCACCGGCGCTCATGACAG 4680
 Db 4621 CTTCCGTAATACCACTCAGGGGAGAGTGTGCTGCTCGCCACCGGCGCTCATGACAG 4680
 Qy 4681 GTATFACTGGGGACTTTGACTCCGTGATCGACTGCAAGCTAGGGGTCACTCAAGTTGTAGA 4740
 Db 4681 GTATFACTGGGGACTTTGACTCCGTGATCGACTGCAAGCTAGGGGTCACTCAAGTTGTAGA 4740

QY 361 TCAAGAAACCAAGAAACCAACACCGTCCGACAGAGCTTTAAAGTTTCCGGCGG 420
 Db 361 TCAAGAAACCAAGAAACCAACACCGTCCGACAGAGCTTTAAAGTTTCCGGCGG 420
 QY 421 CGGCCAGATCGTTGGCGAGTATACCTTTGGCGCGAGGGGCCACAGGTGGGTGGG 480
 Db 421 CGGCCAGATCGTTGGCGAGTATACCTTTGGCGCGAGGGGCCACAGGTGGGTGGG 480
 QY 481 CGCGACAAGAAACACTTCGAGCGGTCCAGCGACCTGGAAGCGCGACCCATCCCTAA 540
 Db 481 CGCGACAAGAAACACTTCGAGCGGTCCAGCGACCTGGAAGCGCGACCCATCCCTAA 540
 QY 541 AGATCGGCGCTCGGCTGGGAGGAACTCTGGGAAACAGGATACCCCTGGGCGGCTTTC 600
 Db 541 AGATCGGCGCTCGGCTGGGAGGAACTCTGGGAAACAGGATACCCCTGGGCGGCTTTC 600
 QY 601 GAATGAGGAGCTCGGCTGGGAGGATGGCTCTGTCGCGGAGGTTCCGCTCCCTTTG 660
 Db 601 GAATGAGGAGCTCGGCTGGGAGGATGGCTCTGTCGCGGAGGTTCCGCTCCCTTTG 660
 QY 661 GGGCCCCAATGACCCCGGATAGGTGCGCAACGCTGGGTAAAGTCAATGATACCTTAAC 720
 Db 661 GGGCCCCAATGACCCCGGATAGGTGCGCAACGCTGGGTAAAGTCAATGATACCTTAAC 720
 QY 721 GTGCGGCTTTGCGGACCTCATGCGGATACATCCCTGCTGGGCGGCGGCTCGGCGG 780
 Db 721 GTGCGGCTTTGCGGACCTCATGCGGATACATCCCTGCTGGGCGGCGGCTCGGCGG 780
 QY 781 CGCGAGAGCTCTCGCGATGCGGTGAGAGTCTGAGAGCGGGGTTAAATTTGCAACAGG 840
 Db 781 CGCGAGAGCTCTCGCGATGCGGTGAGAGTCTGAGAGCGGGGTTAAATTTGCAACAGG 840
 QY 841 GAATTTACCGGTTGCTCTTTTCTATCTTCTGCTGGGCGGCTGCTGCTGCAATACCAC 900
 Db 841 GAATTTACCGGTTGCTCTTTTCTATCTTCTGCTGGGCGGCTGCTGCTGCAATACCAC 900
 QY 901 CCGGCTCTCGGCTCGGAGTGAAGAACATCAGTACCGGCTACATGCTGACTAACGACTG 960
 Db 901 CCGGCTCTCGGCTCGGAGTGAAGAACATCAGTACCGGCTACATGCTGACTAACGACTG 960
 QY 961 CACCAATGACAGATTAACCTGGAGTCAAGGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 Db 961 CACCAATGACAGATTAACCTGGAGTCAAGGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 QY 1021 CCGGCTCGGAGTGGGAGTCACTCAGTCTGAGTACCGGCTCAGCGAGTGGG 1080
 Db 1021 CCGGCTCGGAGTGGGAGTCACTCAGTCTGAGTACCGGCTCAGCGAGTGGG 1080
 QY 1081 CGTGACGGGCGGCGGCTCAGCGAGGCTTGGGAGCGACATCGACATGGTTGTAT 1140
 Db 1081 CGTGACGGGCGGCGGCTCAGCGAGGCTTGGGAGCGACATCGACATGGTTGTAT 1140
 QY 1141 GTGCGCAACGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
 Db 1141 GTGCGCAACGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
 QY 1201 AGCCAAATGTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
 Db 1201 AGCCAAATGTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
 QY 1261 CATCTACCTGGTACCACTGAGACCGCATGGGACATGATGATGATGATGATGATGATGAT 1320
 Db 1261 CATCTACCTGGTACCACTGAGACCGCATGGGACATGATGATGATGATGATGATGATGAT 1320
 QY 1321 GCGCAGGCTACCATGATCTTTGGGATAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
 Db 1321 GCGCAGGCTACCATGATCTTTGGGATAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
 QY 1381 CATTTAGCGGCTCATTTGGGCGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
 Db 1381 CATTTAGCGGCTCATTTGGGCGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
 QY 1441 GTGGCGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500

Db 1441 GTGGCGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
 QY 1501 TGTGCGGAGTCTGCGCGGACACACCGGCGGCTTCAACAGCTTATTTGACATGGGCCC 1560
 Db 1501 TGTGCGGAGTCTGCGCGGACACACCGGCGGCTTCAACAGCTTATTTGACATGGGCCC 1560
 QY 1561 CAGGCGAGAAATCCAGCTCTGTTAAACCAATGGAGCTGGACATCAACCGCACCGCCCT 1620
 Db 1561 CAGGCGAGAAATCCAGCTCTGTTAAACCAATGGAGCTGGACATCAACCGCACCGCCCT 1620
 QY 1621 GAATGCAATGATCTCTTGCACACCGGCTTTATCGGCTCTCTGTTCTACACCGACGCTT 1680
 Db 1621 GAATGCAATGATCTCTTGCACACCGGCTTTATCGGCTCTCTGTTCTACACCGACGCTT 1680
 QY 1681 CAACTGCTCAGGATGTCGGAACGATGTCGCTGCGGAGTATCGAGGCTTCCGGGT 1740
 Db 1681 CAACTGCTCAGGATGTCGGAACGATGTCGCTGCGGAGTATCGAGGCTTCCGGGT 1740
 QY 1741 GGGATGGGCGGCTTGCATATGAGGATATGAGGATATGAGGATATGAGGATATGAGGAT 1800
 Db 1741 GGGATGGGCGGCTTGCATATGAGGATATGAGGATATGAGGATATGAGGATATGAGGAT 1800
 QY 1801 TTGCTGGCACTACCCACCAAGGAGTGTGGCTGCTCGGAGAGCTGTGTGGGCCC 1860
 Db 1801 TTGCTGGCACTACCCACCAAGGAGTGTGGCTGCTCGGAGAGCTGTGTGGGCCC 1860
 QY 1861 AGTGTACTGTTTACCCCGGAGTGTGGCTGCTCGGAGAGCTGTGTGGGCCC 1920
 Db 1861 AGTGTACTGTTTACCCCGGAGTGTGGCTGCTCGGAGAGCTGTGTGGGCCC 1920
 QY 1921 CACTTACAGCTGGGAGGAGATGAGACAGATGCTTCTTCTTATGAAACAGCACTCGACACC 1980
 Db 1921 CACTTACAGCTGGGAGGAGATGAGACAGATGCTTCTTCTTATGAAACAGCACTCGACACC 1980
 QY 1981 GCTGGGCTCATGTTGCGCTGACAGTGAATCTTCTTGGCTACACCAAGATTTGCGG 2040
 Db 1981 GCTGGGCTCATGTTGCGCTGACAGTGAATCTTCTTGGCTACACCAAGATTTGCGG 2040
 QY 2041 CGCACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
 Db 2041 CGCACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
 QY 2101 GGAAGCTTTTGGAGAGCTCTGATACCACTTACCTCAAATGCGGCTCTGGGCGCTGGCT 2160
 Db 2101 GGAAGCTTTTGGAGAGCTCTGATACCACTTACCTCAAATGCGGCTCTGGGCGCTGGCT 2160
 QY 2161 CAGCCCAAGGTGCTGATCGATACCCCTACAGGCTTGGCATTAACCTGCGACAGTTAA 2220
 Db 2161 CAGCCCAAGGTGCTGATCGATACCCCTACAGGCTTGGCATTAACCTGCGACAGTTAA 2220
 QY 2221 CTATACCATCTTCAAATTAAGGATGATGTTGGAGGCTTGGACACAGGCTCAGGCTGTC 2280
 Db 2221 CTATACCATCTTCAAATTAAGGATGATGTTGGAGGCTTGGACACAGGCTCAGGCTGTC 2280
 QY 2281 ATGCAATTTTCACTGCTGGGATCGTTGCAACTTGGAGGACAGAGACAGAACTCAACTGTC 2340
 Db 2281 ATGCAATTTTCACTGCTGGGATCGTTGCAACTTGGAGGACAGAGACAGAACTCAACTGTC 2340
 QY 2341 TCCCTTTTGGCACTCAGCCAGGAAATGGGCAATTTTACCTTGTCTTCTTCTGCGACCTGCC 2400
 Db 2341 TCCCTTTTGGCACTCAGCCAGGAAATGGGCAATTTTACCTTGTCTTCTTCTGCGACCTGCC 2400
 QY 2401 CGCCTTGTGCACTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2460
 Db 2401 CGCCTTGTGCACTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2460
 QY 2461 TGGCCTATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
 Db 2461 TGGCCTATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
 QY 2521 CTTGCTTTAGCGGACCGCAGGTTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580

Ddb	9661	GAGCGCATGACTGCAGAGAGTGCCGTAAC	TCTCTCTCGAGATCATGT	9711
	RESULT 3			
	AAF23486			
ID	AAF23486	standard; DNA; 9711 BP.		
XX	AAF23486;			
XX	21-MAR-2001	(first entry)		
DT	Infectious Hepatitis C virus 2a genotype.			
DE	GBV-B; hepatitis C virus; HCV; vaccine; ds.			
XX	Hepatitis C virus.			
XX	W0200075337-A1.			
PN	14-DEC-2000.			
PD	02-JUN-2000; 2000MO-US15293.			
PP	04-JUN-1999; 99US-0137694.			
PR	(USSH) US DEPT HEALTH & HUMAN SERVICES.			
XX	Bukh J, Yanagi M, Emerson SU, Purcell RH;			
XX	WPI; 2001-091214/10.			
DR	New infectious nucleic acids of the GB virus-B clone, useful for			
DR	indirectly studying the molecular properties of hepatitis C virus			
XX	and in developing vaccines and therapeutics for HCV -			
PS	Disclosure; Page 78-82; 96pp; English.			
XX	The present invention relates to GB virus-B. The nucleic acid mole			
CCC	of the invention are useful for indirectly studying the molecular			
CCC	properties of hepatitis C virus (HCV). The infectious nucleic acid			
CCC	sequence of the GB virus-B clone and the HCV/GBV-B chimeras may be			
CCC	in the development of vaccines and therapeutics for HCV.			
XX	Sequence 9711 BP; 1989 A; 2851 C; 2668 G; 2203 T; 0 other;			
XX	Query Match	100.0%; Score 9707.8; DB 22; Length 9711;		
XX	Best Local Similarity	100.0%; Pred. No. 0;		
XX	Matches 9709; Conservative	0; Mismatches	2; Indels	0; Gaps
Oy	1	ACCGGCCCTAATAGGGGGGACACTCCGCCATGAATCACTCCCTGTGTGAGGAAC	TAC	
Ddb	1	ACCGGCCCTTAATAGGGGGGACACTCCGCCATGAATCACTCCCTGTGTGAGGAAC	TAC	
Oy	61	CTTCACGAGAAAGCGTCTAGCCATGGGTTAGTATAGTCTGCTACAGCCTCCAGAG		
Ddb	61	CTTCACGAGAAAGCGTCTAGCCATGGGTTAGTATAGTCTGCTACAGCCTCCAGAG		
Oy	121	CCCCCTCCCGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCC		
Oy	121	CCCCCTCCCGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCC		
Ddb	181	AAGACTGGGTCTTTCTTGATAAACCACACTCTATGCCCGGCCATTTGGGGCGTCC		
Oy	181	AAGACTGGGTCTTTCTTGATAAACCACACTCTATGCCCGGCCATTTGGGGCGTCC		
Ddb	241	CAAGACTGCTAGCCGAGTAGTCGTTGGGTTGCGAAAGGCGCTTGTGCTACTGCTCAT		
Oy	241	CAAGACTGCTAGCCGAGTAGTCGTTGGGTTGCGAAAGGCGCTTGTGCTACTGCTCAT		
Ddb	301	TGCTTGGAGTGCCCCGGGAGGTCCTGCTAGACCGGTGACACCAATGAGCA	CAAAATCTCTA	
Oy	301	TGCTTGGAGTGCCCCGGGAGGTCCTGCTAGACCGGTGACACCAATGAGCA	CAAAATCTCTA	
Ddb				

QY	6421	CATCATGACCAACGCGTGTCTTGGCGCGCAATATCTCTGGCAATGTTCGGCTTGGGCTC	6480
DB	6421	CATCATGACCAACGCGTGTCTTGGCGCGCAATATCTCTGGCAATGTTCGGCTTGGGCTC	6480
QY	6481	CATGAGATACGCGGCGCTTAAGACCTGCATGAATATCTGCGAGGAGACCTTTCTATCAA	6540
DB	6481	CATGAGATACGCGGCGCTTAAGACCTGCATGAATATCTGCGAGGAGACCTTTCTATCAA	6540
QY	6541	TTGTTTACACGAGGCGCAGTGCCTGCGCGCAAAACCGCGCCAAAACCTTTAAAGTTCGCCATCTG	6600
DB	6541	TTGTTTACACGAGGCGCAGTGCCTGCGCGCAAAACCGCGCCAAAACCTTTAAAGTTCGCCATCTG	6600
QY	6601	GAGGTTGCGGCGCTCAGAGTACGCGGAGGTGACGACGACGCGTCTATACCACTACATAAC	6660
DB	6601	GAGGTTGCGGCGCTCAGAGTACGCGGAGGTGACGACGACGCGTCTATACCACTACATAAC	6660
QY	6661	AGGACTCACCTCATATTAAGTTGAAAGTCCCTGCCCACTACCTCTCCGAGTTCTTTTC	6720
DB	6661	AGGACTCACCTCATATTAAGTTGAAAGTCCCTGCCCACTACCTCTCCGAGTTCTTTTC	6720
QY	6721	CTGGGTGACGAGTGCAGATCCATAGTTTGGCCCCCACACCGAAGCCGTTTTTCCGGGA	6780
DB	6721	CTGGGTGACGAGTGCAGATCCATAGTTTGGCCCCCACACCGAAGCCGTTTTTCCGGGA	6780
QY	6781	TGAGGTCCTGTTGCGGTTGGGCTTAATTTCAATTTGTCGTCGGGTCCAGGCTTCTTTGCGA	6840
DB	6781	TGAGGTCCTGTTGCGGTTGGGCTTAATTTCAATTTGTCGTCGGGTCCAGGCTTCTTTGCGA	6840
QY	6841	CCCTGAACCCGACACAGACGTATTGATGTCTAGCTTAACAGATCCATCTCATATACAGGC	6900
DB	6841	CCCTGAACCCGACACAGACGTATTGATGTCTAGCTTAACAGATCCATCTCATATACAGGC	6900
QY	6901	GGAGACTCAGCGGGCGTTTTAGCGCGGGGTACACCCCATCCGAGCAAGCTCTCCGCG	6960
DB	6901	GGAGACTCAGCGGGCGTTTTAGCGCGGGGTACACCCCATCCGAGCAAGCTCTCCGCG	6960
QY	6961	GAGCGAGCTATCGGCACCATCGCTCGAGGCCACCTGCACCAACCGCAAGCCCTATGA	7020
DB	6961	GAGCGAGCTATCGGCACCATCGCTCGAGGCCACCTGCACCAACCGCAAGCCCTATGA	7020
QY	7021	TGTGACATGTTGATGTAACTGTTTCATGCGGGGGCGATGTGACTCGATAGAGTCTGG	7080
DB	7021	TGTGACATGTTGATGTAACTGTTTCATGCGGGGGCGATGTGACTCGATAGAGTCTGG	7080
QY	7081	GTCCAAAGTGTGTTCTGGAAGTCTCTCGACCCAAATGTCGAAGAAAGGAGCGACCTTGA	7140
DB	7081	GTCCAAAGTGTGTTCTGGAAGTCTCTCGACCCAAATGTCGAAGAAAGGAGCGACCTTGA	7140
QY	7141	GCCTTTCGATACCATCAGAAATACATGCTCCCCAAGAAGAGTTCCTCACAGCTTTTACCGGC	7200
DB	7141	GCCTTTCGATACCATCAGAAATACATGCTCCCCAAGAAGAGTTCCTCACAGCTTTTACCGGC	7200
QY	7201	CTGGGCAACGCGCTGATTAACAACCAACCGCTTGTGGAATCTGTGGAAGAGCGCAGATTACCA	7260
DB	7201	CTGGGCAACGCGCTGATTAACAACCAACCGCTTGTGGAATCTGTGGAAGAGCGCAGATTACCA	7260
QY	7261	ACCGGCGACTGTTGGGCGTGTCTCTCCCTCTCTAGGAAACCCCGAGCGCTCCCC	7320
DB	7261	ACCGGCGACTGTTGGGCGTGTCTCTCCCTCTCTAGGAAACCCCGAGCGCTCCCC	7320
QY	7321	AAGGAGGCGCGGACAGTGGGCTTAAGTGAAGACTCCATAGGAGATGCCCTTCAAACAGCT	7380
DB	7321	AAGGAGGCGCGGACAGTGGGCTTAAGTGAAGACTCCATAGGAGATGCCCTTCAAACAGCT	7380
QY	7381	GGCCATTAAAGTCTTTGGCCAGCGCCCCCAAGCGGCGATTACGCGCTTTTCCACGGGGC	7440
DB	7381	GGCCATTAAAGTCTTTGGCCAGCGCCCCCAAGCGGCGATTACGCGCTTTTCCACGGGGC	7440
QY	7441	GGGCGCTGCGGATTCGGGCGATCAGACGCTCTCTGATGAGTTGGCCCTTTTCGAGACAGG	7500
DB	7441	GGGCGCTGCGGATTCGGGCGATCAGACGCTCTCTGATGAGTTGGCCCTTTTCGAGACAGG	7500

Qy	2041	CGCACACCTGCGCTAGCTAGCTGACTTCAACGCCAGCAGCAGCAGCTGTTGCCCCAC	2100	Qy	3121	CATGGCAAGGCACTCTCGCGGGGGGAGGTACTCTCAGATGGCGCTACTAGCCCTTGGCAG	3180
Db	2041	CGCACACCTGCGCTAGCTAGCTGACTTCAACGCCAGCAGCAGCAGCTGTTGCCCCAC	2100	Db	3121	CATGGCAAGGCACTCTCGCGGGGGGAGGTACTCTCAGATGGCGCTACTAGCCCTTGGCAG	3180
Qy	2101	GGACTGTTTGGAAAGCATCTGATACCACTTACCTCAAAATGGCGCTCTGGGCCCTGGCT	2160	Qy	3181	GTGGACTGGCACTTACATCTATGACCACTGACCCCTATGTGGATTTGGGCTGTAGTGG	3240
Db	2101	GGACTGTTTGGAAAGCATCTGATACCACTTACCTCAAAATGGCGCTCTGGGCCCTGGCT	2160	Db	3181	GTGGACTGGCACTTACATCTATGACCACTGACCCCTATGTGGATTTGGGCTGTAGTGG	3240
Qy	2161	CAGCCNAGGTGCTGATGACCTACCCCTACAGCTCTGCAATACCCCTGACAGTTAA	2220	Qy	3241	CCTGGGGGACCTGGCGGTGCGCTTGGAGCTTATCATCTTTCAGTCCGATGGAGAGAAAGT	3300
Db	2161	CAGCCNAGGTGCTGATGACCTACCCCTACAGCTCTGCAATACCCCTGACAGTTAA	2220	Db	3241	CCTGGGGGACCTGGCGGTGCGCTTGGAGCTTATCATCTTTCAGTCCGATGGAGAGAAAGT	3300
Qy	2221	CTATACCATCTTCAAAATAGGATGTATGTGGAGGGGTTGAGCAGAGGCTCAGCGCTGC	2280	Qy	3301	CATTTCTGGGAGCGGAGACAGCTGCTTGTGGGAGACATTTTACAGCACTTCCCTGTGC	3360
Db	2221	CTATACCATCTTCAAAATAGGATGTATGTGGAGGGGTTGAGCAGAGGCTCAGCGCTGC	2280	Db	3301	CATTTCTGGGAGCGGAGACAGCTGCTTGTGGGAGACATTTTACAGCACTTCCCTGTGC	3360
Qy	2281	ATGCAATTTCACTCGTGGGATCGTTGCAATTTGGAGGACAGAGACAGAAAGTCAACTGTC	2340	Qy	3361	CGCCGACCTTGTGGGAGGCTCTCTTGGCCCGAGCTGATGCTATACCTTCCAGGGGTG	3420
Db	2281	ATGCAATTTCACTCGTGGGATCGTTGCAATTTGGAGGACAGAGACAGAAAGTCAACTGTC	2340	Db	3361	CGCCGACCTTGTGGGAGGCTCTCTTGGCCCGAGCTGATGCTATACCTTCCAGGGGTG	3420
Qy	2341	TCCTTTGTCGACTCCACACGGAATGGGCCATTTTACCTTGTCTTCTTCTCGGACCTGCC	2400	Qy	3421	GAGTCTTCTCGCCCCCATCACTGCTTACGCCCGAGCAGACACGTTGGGCACTT	3480
Db	2341	TCCTTTGTCGACTCCACACGGAATGGGCCATTTTACCTTGTCTTCTTCTCGGACCTGCC	2400	Db	3421	GAGTCTTCTCGCCCCCATCACTGCTTACGCCCGAGCAGACACGTTGGGCACTT	3480
Qy	2401	CGCCTTGTGCACTGCTGCGGATCGTTGCAATTTGGAGGACAGAGACAGAAAGTCAACTGTA	2460	Qy	3481	AGTGTGAGCATGACGCGGGCGGACAGACAGACAGGCTGGGAAATTCAGTCTCTGTC	3540
Db	2401	CGCCTTGTGCACTGCTGCGGATCGTTGCAATTTGGAGGACAGAGACAGAAAGTCAACTGTA	2460	Db	3481	AGTGTGAGCATGACGCGGGCGGACAGACAGACAGGCTGGGAAATTCAGTCTCTGTC	3540
Qy	2461	TGGGCTATACCTGCTCCCAAAATACATCGTGGAGCGTACAAATTCATGTA	2520	Qy	3541	CACAGTCACTCAGTCTCTCTCGGAACTCTCTCGGGGGTCTTGGGAGCTGTCTACCA	3600
Db	2461	TGGGCTATACCTGCTCCCAAAATACATCGTGGAGCGTACAAATTCATGTA	2520	Db	3541	CACAGTCACTCAGTCTCTCTCGGAACTCTCTCGGGGGTCTTGGGAGCTGTCTACCA	3600
Qy	2521	CCTGCTTTAGCGGACGCGAGGTTTGGGCTGCTTATGATGCTCATCTTGTGGGCCA	2580	Qy	3601	TGGAGCTGGCAACAAGACTCTGGCCGCTCAGGGGTCGGGTCCGGTCAAGATGTACTCCAG	3660
Db	2521	CCTGCTTTAGCGGACGCGAGGTTTGGGCTGCTTATGATGCTCATCTTGTGGGCCA	2580	Db	3601	TGGAGCTGGCAACAAGACTCTGGCCGCTCAGGGGTCGGGTCCGGTCAAGATGTACTCCAG	3660
Qy	2581	GGCCGAGCAGCACTAGAGAGCTGCTGCTTCTGACGCTGCGAGCGCAGCTAGCTGCA	2640	Qy	3661	TGCTGAGGGGACTTATAGTGGGTGGCCAGCCCTCTGGGACTAAATCTTGGAGCGCTG	3720
Db	2581	GGCCGAGCAGCACTAGAGAGCTGCTGCTTCTGACGCTGCGAGCGCAGCTAGCTGCA	2640	Db	3661	TGCTGAGGGGACTTATAGTGGGTGGCCAGCCCTCTGGGACTAAATCTTGGAGCGCTG	3720
Qy	2641	TGGCTTCTATATTTTGTGCTATCTTTTGTGGCTGCTTGTGATCATCAAGGGTGGGTAGT	2700	Qy	3721	CAGTGTGAGCGGTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3780
Db	2641	TGGCTTCTATATTTTGTGCTATCTTTTGTGGCTGCTTGTGATCATCAAGGGTGGGTAGT	2700	Db	3721	CAGTGTGAGCGGTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3780
Qy	2701	CCCCCTAGCTACCTATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2760	Qy	3781	AAGACCGGGGCAAAACGGGAGCGCTACTCTCTCCCGAGACCTCTTTCACCTTTGAAGGG	3840
Db	2701	CCCCCTAGCTACCTATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2760	Db	3781	AAGACCGGGGCAAAACGGGAGCGCTACTCTCTCCCGAGACCTCTTTCACCTTTGAAGGG	3840
Qy	2761	GGCCCAACAGGCTTATGCTTATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2820	Qy	3841	GTCTCAGGAGGCGCGGTGCTATGCGCCAGGGGCGCAGCTGTGCGAGTCTTCCGGGCGC	3900
Db	2761	GGCCCAACAGGCTTATGCTTATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2820	Db	3841	GTCTCAGGAGGCGCGGTGCTATGCGCCAGGGGCGCAGCTGTGCGAGTCTTCCGGGCGC	3900
Qy	2821	GGTAAATGATCACTCTCTTACTCTACCCCGGGTATAAGACCTTCTCAGCGGGTTTTT	2880	Qy	3901	TGTTGCTCTCGGGGCGTGAAGTCAATAGATTTTATCCCGCTTGGAGACTTCCGACAT	3960
Db	2821	GGTAAATGATCACTCTCTTACTCTACCCCGGGTATAAGACCTTCTCAGCGGGTTTTT	2880	Db	3901	TGTTGCTCTCGGGGCGTGAAGTCAATAGATTTTATCCCGCTTGGAGACTTCCGACAT	3960
Qy	2881	GTGGTGTGCTGCTATCTTCTGACCTTGGGGAGCTATGCTGCGAGGTGGGACCCACC	2940	Qy	3961	CGTCAGCGGCTCCCCCACTTTTGTAGTGAACAACAGCAGCAGCAGCAGCAGCAGCAGC	4020
Db	2881	GTGGTGTGCTGCTATCTTCTGACCTTGGGGAGCTATGCTGCGAGGTGGGACCCACC	2940	Db	3961	CGTCAGCGGCTCCCCCACTTTTGTAGTGAACAACAGCAGCAGCAGCAGCAGCAGCAGC	4020
Qy	2941	TATGAGGTGGCGGCTGCTGATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3000	Qy	4021	TCAGGTGGGTACTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	4080
Db	2941	TATGAGGTGGCGGCTGCTGATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3000	Db	4021	TCAGGTGGGTACTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	4080
Qy	3001	TGCTGTTTGAATAACCAAGTGGCTTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3060	Qy	4081	ATATGCTGCTCAGGGGTATAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT	4140
Db	3001	TGCTGTTTGAATAACCAAGTGGCTTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3060	Db	4081	ATATGCTGCTCAGGGGTATAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT	4140
Qy	3061	AGTGTCTTGAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3120	Qy	4141	GTTTGGGGCTACTTGTCTAAGGCAATGTCATCAATCCCAACATTTAGGACTGTGAGTGTG	4200
Db	3061	AGTGTCTTGAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3120	Db	4141	GTTTGGGGCTACTTGTCTAAGGCAATGTCATCAATCCCAACATTTAGGACTGTGAGTGTG	4200
				Qy	4201	GACTGTGAAGCGGCGGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	4260

8341	CTCAAACCGTCACTGAGAGAGACATCAGGACTCAGGAGTCCATATATATCGGGCCCTGCTCCTT	QY
8341	CTCAAACCGTCACTGAGAGAGACATCAGGACTCAGGAGTCCATATATATCGGGCCCTGCTCCTT	Db
8401	GCCCGAGGAGGCCACACACTGCCATACACTCGCTAACTGAGAGACATTTACGTGGGAGGGCC	QY
8401	GCCCGAGGAGGCCACACATGCCATATACATCTGCTTACTGAGAGACTTTACGTGGGAGGGCC	Db
8461	TATGTTTCAACAGCAAGGGGCCAAACCTCGGGGTACAGGGGTTCGCCGCCAGCGGGGTGCT	QY
8461	TATGTTTCAACAGCAAGGGGCCAAACCTCGGGGTACAGGGGTTCGCCGCCAGCGGGGTGCT	Db
8521	CACCACTAGCATGGGGAACACCATCACATGCTACGTGAAGGCTTTAGCGGCTTTGTAAGC	QY
8521	CACCACTAGCATGGGGAACACCATCACATGCTACGTGAAGGCTTTAGCGGCTTTGTAAGC	Db
8581	TGCAAGGATAATCGCGGCCACAAATGCTGTATGCGCGGATGACTTGGTTGTTCATCTCAGA	QY
8581	TGCAAGGATAATCGCGGCCACAAATGCTGTATGCGCGGATGACTTGGTTGTTCATCTCAGA	Db
8641	AAGCCAGGGGACCGAGGAGGACGAGCGGAACCTGAGAGCCTTACCGAGGCTATGACCCAG	QY
8641	AAGCCAGGGGACCGAGGAGGACGAGCGGAACCTGAGAGCCTTACCGAGGCTATGACCCAG	Db
8701	GTATTCTGCCCTCTCTGTGTACCCCCCHAGACCGGAGTATGATCTGGAGCTGATACATC	QY
8701	GTATTCTGCCCTCTCTGTGTACCCCCCHAGACCGGAGTATGATCTGGAGCTGATACATC	Db
8761	TTGCTCTCAAATGTCTGTGGCGCTGGGCCCCACAAGGCCCGCGCAGATATCTACTCTGAC	QY
8761	TTGCTCTCAAATGTCTGTGGCGCTGGGCCCCACAAGGCCCGCGCAGATATCTACTCTGAC	Db
8821	CAGAGACCTTACCACTCCAATCGCCGGGCTGCCTGGGAAAACAGTTTAGACATCTCCCTGT	QY
8821	CAGAGACCTTACCACTCCAATCGCCGGGCTGCCTGGGAAAACAGTTTAGACATCTCCCTGT	Db
8881	CAATTCATGGCTGGGAAAACATCATCCAGTACGCCCGCGACCATATGGGCTCGCATGGTCT	QY
8881	CAATTCATGGCTGGGAAAACATCATCCAGTACGCCCGCGACCATATGGGCTCGCATGGTCT	Db
8941	GATGACACACTTCTTCTCCATTTCTCATGTCTCAAGACACGCTGGACAGAACTCAACTT	QY
8941	GATGACACACTTCTTCTCCATTTCTCATGTCTCAAGACACGCTGGACAGAACTCAACTT	Db
9001	TGAGATGTACGAGCGGTGTACTCCGTGAGTCCCTTGGACCTCCAGCTATTAATTGAAAG	QY
9001	TGAGATGTACGAGCGGTGTACTCCGTGAGTCCCTTGGACCTCCAGCTATTAATTGAAAG	Db
9061	GTTACATGGGCTTGACGCTTTTCTCTGCACACATACATCTCCGACGAATGACACGGGT	QY
9061	GTTACATGGGCTTGACGCTTTTCTCTGCACACATACATCTCCGACGAATGACACGGGT	Db
9121	GGCTTCAGCCCTCAGAAAAACTTGGGGCGCCACCCCTCAGAGCGTGAAGAGCGGGGACG	QY
9121	GGCTTCAGCCCTCAGAAAAACTTGGGGCGCCACCCCTCAGAGCGTGAAGAGCGGGGACG	Db
9181	TGCAGTCAAGGGGTCCCTCATCTCCGTGGGGGAGAGCGGCGTTTGGCGTGCATATCT	QY
9181	TGCAGTCAAGGGGTCCCTCATCTCCGTGGGGGAGAGCGGCGTTTGGCGTGCATATCT	Db
9241	CTTCAATGGGCGGTGAAGACCAAGCTCAAACTCACTCCATTCGCCGGAAGCGGCTCCT	QY
9241	CTTCAATGGGCGGTGAAGACCAAGCTCAAACTCACTCCATTCGCCGGAAGCGGCTCCT	Db
9301	GGATTTATCCAGCTGGTTCAACCGTCGGCGCGCGGGGGCGACATTTTATCACAGCGTGT	QY
9301	GGATTTATCCAGCTGGTTCAACCGTCGGCGCGCGGGGGCGACATTTTATCACAGCGTGT	Db
9361	GGGTGCCGACCCCGCTTATGTCTTTGGCCTPACTCTCTATCTTTTGTAGGGGTAGGCT	QY
9361	GGGTGCCGACCCCGCTTATGTCTTTGGCCTPACTCTCTATCTTTTGTAGGGGTAGGCT	Db

	QY	9421	TTTTCTACTCCCGCTCGGTAGAGGGGCACACATTTAGCTTACACTCCTTAGCTAACTGTC	9480
	Db	9421	TTTTCTACTCCCGCTCGGTAGAGGGGCACACATTTAGCTTACACTCCTTAGCTAACTGTC	9480
	QY	9481	CTT	9540
	Db	9481	CTT	9540
	QY	9541	TTT	9600
	Db	9541	TTT	9600
	QY	9601	TACTTTCCTTGCTGGTGGCTCCATCTTAGCCCTAGTCA CGGCTAGCTGGA AGGTC CGT	9660
	Db	9601	TACTTTCCTTGCTGGTGGCTCCATCTTAGCCCTAGTCA CGGCTAGCTGGA AGGTC CGT	9660
	QY	9661	GAGCGCATGACTGCAGAGAGTGC CGTA ACTGGTCTCTCTGCAGATCATGT	9711
	Db	9661	GAGCGCATGACTGCAGAGAGTGC CGTA ACTGGTCTCTCTGCAGATCATGT	9711
RESULT 2				
	ID	AAC86937	AAC86937 standard; DNA; 9711 BP.	
	XX	XX	AAC86937;	
	XX	XX	02-APR-2001 (first entry)	
	XX	XX	Nucleotide sequence of a hepatitis C virus (HCV) clone genotype 2a.	
	DE	XX	Chimeric virus; bovine viral diarrhoea virus; BVDV; hepatitis C virus;	
	KW	XX	HCV; vaccine; viral inhibitor; antiviral; ss.	
	KX	XX	Hepatitis C virus.	
	OS	XX	Key Location/Qualifiers	
	FH	XX	CDS 341..9442	
	FT	XX	/*tag= a	
	FT	XX	WO200075352-A2.	
	PB	XX	14-DEC-2000.	
	PF	XX	02-JUN-2000; 2000WO-US15527.	
	PP	XX	04-JUN-1999; 99US-0137817.	
	PR	XX	(USSH) US DEPT HEALTH & HUMAN SERVICES.	
	PA	XX	Nam J, Bukh J, Emerson SU, Purcell RH;	
	PI	XX	WPI; 2001-071081/08.	
	DR	XX	P-PSDB; AAB31168.	
	DR	XX	New nucleic acid comprising a chimeric bovine viral diarrhoea virus genome in which the (non-) structural region has been replaced by hepatitis C virus (HCV) genome useful for treating or preventing HCV signs and symptoms -	
	PT	XX	Disclosure; Page 82-85; 97pp; English.	
	PT	XX	The specification describes a nucleic acid comprising a chimeric virus genome, specifically bovine viral diarrhoea virus (BVDV) genome in which the (non-)structural region has been replaced by the (non-)structural region of a hepatitis C virus (HCV) genome. The nucleic acids comprising the chimeric virus and the chimeric virus are useful for identifying cell lines capable of supporting the replication of these chimeric viruses, in screening for neutralizing antibodies to HCV of different genotypes, in the production of HCV-BVDV virions, for the development of inactivated or attenuated vaccines to prevent HCV-BVDV in a mammal, in studying the molecular properties of HCV indirectly in vitro, and in identifying inhibitors of viral enzyme activity which would be useful	

Db	6121	TTCCAGAGAAATCACTGCGCCCAACCCACTACGTACGTGACGGAGTCGGATCGCTGCGAGCG	6180	Db	7201	CTGGCAGCGGCTGATTAACAACCCACCGCTTGTGGAATCGTGGAAAAGCCAGATTACCA	7260
Qy	6181	TGTGACCCAACTACTTGGCTCCCTTACCATAACACCGCTGCTCAGAAAGACTCCAACTG	6240	Qy	7261	ACCGCCCACTGTTGGGGCTGTCTCTCCCTCTAGGAAAACCCCGAGCGCTCCGCC	7320
Db	6181	TGTGACCCAACTACTTGGCTCCCTTACCATAACACCGCTGCTCAGAAAGACTCCAACTG	6240	Db	7261	ACCGCCCACTGTTGGGGCTGTCTCTCCCTCTAGGAAAACCCCGAGCGCTCCGCC	7320
Qy	6241	GATTACTGAGGACTGCCCATCCATCCATCGCGGCTCGTGGCTCCGCGATGTGGGACTG	6300	Qy	7321	TAAGAGGCGCGGACAGTGGGCTTAAGTGAGGACTCCATAGGAGATGCCCTTCAACAGCT	7380
Db	6241	GATTACTGAGGACTGCCCATCCATCCATCGCGGCTCGTGGCTCCGCGATGTGGGACTG	6300	Db	7321	TAAGAGGCGCGGACAGTGGGCTTAAGTGAGGACTCCATAGGAGATGCCCTTCAACAGCT	7380
Qy	6301	GCTTTCACCATCTTAACAGACTTTAAATAATGGCTGACCTCCAAATATATCCCAAGAT	6360	Qy	7381	GGCCATTAAGTCTTTGGCCAGCCGCCCAAGCGCGATTTCAGGCTTTTCCACGGGCG	7440
Db	6301	GCTTTCACCATCTTAACAGACTTTAAATAATGGCTGACCTCCAAATATATCCCAAGAT	6360	Db	7381	GGCCATTAAGTCTTTGGCCAGCCGCCCAAGCGCGATTTCAGGCTTTTCCACGGGCG	7440
Qy	6361	GCCCGGCTCCCTTGTCTCTGTCAAAAGGGTACAAGGGGCTGTGGCGCGCACTGG	6420	Qy	7441	GGCGCTCCGATTCGCGAGTCAGACGCTCTCTGATGATGGGCTTTTCGGAGACAGG	7500
Db	6361	GCCCGGCTCCCTTGTCTCTGTCAAAAGGGTACAAGGGGCTGTGGCGCGCACTGG	6420	Db	7441	GGCGCTCCGATTCGCGAGTCAGACGCTCTCTGATGATGGGCTTTTCGGAGACAGG	7500
Qy	6421	CATCATGACACACGCTGCTTGGCGGCAATATCTTGGCAATGTCCGCTTGGGCTC	6480	Qy	7501	TTCCATCTTCTTCATGCGCCCTCGAGGGGAGCTTGGAGATCCAGACCTGGAGCGCTGA	7560
Db	6421	CATCATGACACACGCTGCTTGGCGGCAATATCTTGGCAATGTCCGCTTGGGCTC	6480	Db	7501	TTCCATCTTCTTCATGCGCCCTCGAGGGGAGCTTGGAGATCCAGACCTGGAGCGCTGA	7560
Qy	6481	CATGAGAAATCAAGGCTTAAGACCTGCAATGATATCTGGCAGGAGACCTTTCTATCAA	6540	Qy	7561	GCAGGTAGAGCCCAACCCCGCCAGGGGGGGTGGCAGCTCCCGGCTCGGACTCGGG	7620
Db	6481	CATGAGAAATCAAGGCTTAAGACCTGCAATGATATCTGGCAGGAGACCTTTCTATCAA	6540	Db	7561	GCAGGTAGAGCCCAACCCCGCCAGGGGGGGTGGCAGCTCCCGGCTCGGACTCGGG	7620
Qy	6541	TTGTTACACGGAGGCGAGTGGCGGCAACCCCGCCAGGCTTTAAGGTCGCGCATCTG	6600	Qy	7621	GTCTGTGCTACTTCTCGGAGGAGCGACTCCGCTCGTGTGCTCTCATGTCTACTCTC	7680
Db	6541	TTGTTACACGGAGGCGAGTGGCGGCAACCCCGCCAGGCTTTAAGGTCGCGCATCTG	6600	Db	7621	GTCTGTGCTACTTCTCGGAGGAGCGACTCCGCTCGTGTGCTCTCATGTCTACTCTC	7680
Qy	6601	GAGGTTGGGCTCAGAGTACCGGAGGTGACGAGACGCGTCAATACCTACATAAC	6660	Qy	7681	CTGGACCGGGCTCTAATAACTCTTGTAGTCCCGAGAGGAGAGTGTACCGATTAAACC	7740
Db	6601	GAGGTTGGGCTCAGAGTACCGGAGGTGACGAGACGCGTCAATACCTACATAAC	6660	Db	7681	CTGGACCGGGCTCTAATAACTCTTGTAGTCCCGAGAGGAGAGTGTACCGATTAAACC	7740
Qy	6661	AGGACTCACCACATGATTAAGTTCGCTGCACTACCTCTCCGAGTTCCTTTTC	6720	Qy	7741	CTTGAGCAACTCCCTGTTGCGATATCAACAAGGTGTACTGTACCAACAAGAGGCG	7800
Db	6661	AGGACTCACCACATGATTAAGTTCGCTGCACTACCTCTCCGAGTTCCTTTTC	6720	Db	7741	CTTGAGCAACTCCCTGTTGCGATATCAACAAGGTGTACTGTACCAACAAGAGGCG	7800
Qy	6721	CTGGTGAGCGAGTGCAGATCCATAGTTTGGCCCGCCACACGAGCGGTTTTTCGGGA	6780	Qy	7801	CTCAGTAAAGGCTAAGAGTAACTTTGATAGGATGCAAGTGTCTCTACTACGA	7860
Db	6721	CTGGTGAGCGAGTGCAGATCCATAGTTTGGCCCGCCACACGAGCGGTTTTTCGGGA	6780	Db	7801	CTCAGTAAAGGCTAAGAGTAACTTTGATAGGATGCAAGTGTCTCTACTACGA	7860
Qy	6781	TGAGGCTCTGCTGCTGGGCTTAATCATTTGCTGCGGTCCAGCTTCCCTTTCGGA	6840	Qy	7861	CTCAGTCTTAAAGGACATTTAAGTAGCGGCTTCAAGGTCAAGGTCTCTCACCAT	7920
Db	6781	TGAGGCTCTGCTGCTGGGCTTAATCATTTGCTGCGGTCCAGCTTCCCTTTCGGA	6840	Db	7861	CTCAGTCTTAAAGGACATTTAAGTAGCGGCTTCAAGGTCAAGGTCTCTCACCAT	7920
Qy	6841	CCCTGAACCCGACACAGAGCTATTGATGTCATGCTTAAAGATCCATCTCATATCAGCG	6900	Qy	7921	GGAGGAGCTTGCAGTTAAACCCCACTTCTCAAGATCTAAATATGGGTTTGGGCG	7980
Db	6841	CCCTGAACCCGACACAGAGCTATTGATGTCATGCTTAAAGATCCATCTCATATCAGCG	6900	Db	7921	GGAGGAGCTTGCAGTTAAACCCCACTTCTCAAGATCTAAATATGGGTTTGGGCG	7980
Qy	6901	GGAGCTGACGCGGCTTTAGCGGGGGTCAACCCCATTCGAGGCAAGTCTCTCGG	6960	Qy	7981	TAAGAGGCTCCGAGCTTGTCCGGAGGGGTTAACCAATCAAGTCCGTTGGGAAGA	8040
Db	6901	GGAGCTGACGCGGCTTTAGCGGGGGTCAACCCCATTCGAGGCAAGTCTCTCGG	6960	Db	7981	TAAGAGGCTCCGAGCTTGTCCGGAGGGGTTAACCAATCAAGTCCGTTGGGAAGA	8040
Qy	6961	GAGCAGCTATCGGACCATCGCTGAGCCACCTGACACACCGAGGAGGCTTATGA	7020	Qy	8041	CCTCTCGAGGACTCAGAAACCAATTTCCCAACACCATTTATGGCCAAAATGAGGTGT	8100
Db	6961	GAGCAGCTATCGGACCATCGCTGAGCCACCTGACACACCGAGGAGGCTTATGA	7020	Db	8041	CCTCTCGAGGACTCAGAAACCAATTTCCCAACACCATTTATGGCCAAAATGAGGTGT	8100
Qy	7021	TGTGGACATGGTGGATGCTAACTGTTTATGGGGGGGATGTGACTCCGATAGGCTCG	7080	Qy	8101	CTGCGTGACCCCAACAGGGGGGCAAGAAAGCAGCTCGCTTATCGTTTACCTGACCT	8160
Db	7021	TGTGGACATGGTGGATGCTAACTGTTTATGGGGGGGATGTGACTCCGATAGGCTCG	7080	Db	8101	CTGCGTGACCCCAACAGGGGGGCAAGAAAGCAGCTCGCTTATCGTTTACCTGACCT	8160
Qy	7081	GTCCAAAGTGTGCTTGTGACTCTCTCGACCCCAATGGTTCGAAGAAAGGAGGACCTTGA	7140	Qy	8161	CGGCGTCAAGGCTCTCGAGAGAGTGGCTTTATGACATTAACAACCTCTCTCAGC	8220
Db	7081	GTCCAAAGTGTGCTTGTGACTCTCTCGACCCCAATGGTTCGAAGAAAGGAGGACCTTGA	7140	Db	8161	CGGCGTCAAGGCTCTCGAGAGAGTGGCTTTATGACATTAACAACCTCTCTCAGC	8220
Qy	7141	GCCTTCGATACATCAGATACATGCTCCCAAGAGAGGTTCCCAACAGCTTACCGGC	7200	Qy	8221	GGTGTAGGGGCTCTTATGATTCAGTATTTCCCGCTCAGCGGGTAGAGTTCTCTT	8280
Db	7141	GCCTTCGATACATCAGATACATGCTCCCAAGAGAGGTTCCCAACAGCTTACCGGC	7200	Db	8221	GGTGTAGGGGCTCTTATGATTCAGTATTTCCCGCTCAGCGGGTAGAGTTCTCTT	8280
Qy	7201	CTGGCAGCGGCTGATTAACAACCCCGCTTGTGGAAATCGTGGAAAAGCCAGATTACCA	7260	Qy	8281	GAAAGCATGGCGGAAAAGAGGACCTTATGGGTTTTTCTGATGATACCCGATGCTTGA	8340
Db	7201	CTGGCAGCGGCTGATTAACAACCCCGCTTGTGGAAATCGTGGAAAAGCCAGATTACCA	7260	Db	8281	GAAAGCATGGCGGAAAAGAGGACCTTATGGGTTTTTCTGATGATACCCGATGCTTGA	8340

1741	GGGATGGGGCGCTTGCAATATGAGGATAATGTCAACCAATCCAGAGATATGAGACCCCTA	1800	2821	GGTAATGATCACTCTCTTTACTCTCACCCCGGGTATAAGACCTTTCTCAGCGGTTTTT	2880
1801	TTGCTGGCACTACCCACCAAGGAGTGTGGCGTGTCTCGGCGAAGACTGTGTGGCCC	1860	2881	GTGGTGTGTGCTATCTTTGACCTTGGGGAGAGCTATGGTCCAGGAGTGGGCAACCC	2940
1801	TTGCTGGCACTACCCACCAAGGAGTGTGGCGTGTCTCGGCGAAGACTGTGTGGCCC	1860	2881	GTGGTGTGTGCTATCTTTGACCTTGGGGAGAGCTATGGTCCAGGAGTGGGCAACCC	2940
1861	AGTGTACTCTTTCACCCCGGAGGAGTGTGGCGTGTCTCGGCGAAGACTGTGTGGCCC	1920	2941	TATGAGGTGGCGGCTGTGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCGTGTCT	3000
1861	AGTGTACTCTTTCACCCCGGAGGAGTGTGGCGTGTCTCGGCGAAGACTGTGTGGCCC	1920	2941	TATGAGGTGGCGGCTGTGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCGTGTCT	3000
1921	CACCTACAGTGGGGGAGGAGTGTGGCGTGTCTCGGCGAAGACTGTGTGGCCC	1980	3001	TGTGTGTGTGACATTAACCAAGTGTGTGGCGTGTCTCGGCGTGTCTCGGCGTGTCT	3060
1921	CACCTACAGTGGGGGAGGAGTGTGGCGTGTCTCGGCGAAGACTGTGTGGCCC	1980	3001	TGTGTGTGTGACATTAACCAAGTGTGTGGCGTGTCTCGGCGTGTCTCGGCGTGTCT	3060
1981	GCTGGGTGATGTTGGGCTGAGTGGTGTGGCGTGTCTCGGCGAAGACTGTGTGGCCC	2040	3061	AGTGTGTGTGACATTAACCAAGTGTGTGGCGTGTCTCGGCGTGTCTCGGCGTGTCT	3120
1981	GCTGGGTGATGTTGGGCTGAGTGGTGTGGCGTGTCTCGGCGAAGACTGTGTGGCCC	2040	3061	AGTGTGTGTGACATTAACCAAGTGTGTGGCGTGTCTCGGCGTGTCTCGGCGTGTCT	3120
2041	CGCACACCTGCGCTAGTGTGGTGTGGCGTGTCTCGGCGAAGACTGTGTGGCCC	2100	3121	CATGGCAAGGCACTCTCGGCGGCGGAGTATCTCGGCGTGTCTCGGCGTGTCTCGGCG	3180
2041	CGCACACCTGCGCTAGTGTGGTGTGGCGTGTCTCGGCGAAGACTGTGTGGCCC	2100	3121	CATGGCAAGGCACTCTCGGCGGCGGAGTATCTCGGCGTGTCTCGGCGTGTCTCGGCG	3180
2101	GGACTGTTTGGGAGGATCTCTGATACCACTTACCTCAATGCGGCTCTGGGCGCTGGCT	2160	3181	GTGGACTGGCACTTATCATCTATGACCACTCAACCTTATGTGGGCTGTCTCGGCGT	3240
2101	GGACTGTTTGGGAGGATCTCTGATACCACTTACCTCAATGCGGCTCTGGGCGCTGGCT	2160	3181	GTGGACTGGCACTTATCATCTATGACCACTCAACCTTATGTGGGCTGTCTCGGCGT	3240
2161	CAGCCAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2220	3241	CCTGGGAGGAGTGTGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCG	3300
2161	CAGCCAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2220	3241	CCTGGGAGGAGTGTGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCG	3300
2221	CTATACATCTTCAAAATAGGATGATGATGATGATGATGATGATGATGATGATGATGAT	2280	3301	CATTTGTGGGAGGAGTGTGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCGTGTCT	3360
2221	CTATACATCTTCAAAATAGGATGATGATGATGATGATGATGATGATGATGATGATGAT	2280	3301	CATTTGTGGGAGGAGTGTGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCGTGTCT	3360
2281	ATGCAATTTTCACTCGTGGGATGATGATGATGATGATGATGATGATGATGATGATGAT	2340	3361	CGCCGAGTGTGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCGTGTCT	3420
2281	ATGCAATTTTCACTCGTGGGATGATGATGATGATGATGATGATGATGATGATGATGAT	2340	3361	CGCCGAGTGTGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCGTGTCT	3420
2341	TCCTTTGTGGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2400	3421	GAGTCTTCTCGGCGGAGTGTGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCGTGTCT	3480
2341	TCCTTTGTGGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2400	3421	GAGTCTTCTCGGCGGAGTGTGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCGTGTCT	3480
2401	CGCTTGTGGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2460	3481	AGTGTGTAGCATGAGCGGCGGAGTGTGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCG	3540
2401	CGCTTGTGGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2460	3481	AGTGTGTAGCATGAGCGGCGGAGTGTGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCG	3540
2461	TGGCTTATCACTGCTTCAAAATACATGCTCGATGGGAGTGTGGCGTGTCTCGGCGTGTCT	2520	3541	CACAGTCTAGTCTTCTCGGCGGAGTGTGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCG	3600
2461	TGGCTTATCACTGCTTCAAAATACATGCTCGATGGGAGTGTGGCGTGTCTCGGCGTGTCT	2520	3541	CACAGTCTAGTCTTCTCGGCGGAGTGTGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCG	3600
2521	CCTGCTTTAGCGGAGCGGAGGTTTGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCGTGTCT	2580	3601	TGGAGTGGCAACAGACTCTGGCGGAGTGTGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCG	3660
2521	CCTGCTTTAGCGGAGCGGAGGTTTGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCGTGTCT	2580	3601	TGGAGTGGCAACAGACTCTGGCGGAGTGTGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCG	3660
2581	GGCGGAGGAGTGTGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCGTGTCT	2640	3661	TGCTGAGGGGAGTGTGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCGTGTCT	3720
2581	GGCGGAGGAGTGTGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCGTGTCT	2640	3661	TGCTGAGGGGAGTGTGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCGTGTCT	3720
2641	TGGCTTCTATATTTTGTGATCTTTTGTGCTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT	2700	3721	CACGTGTGGAGCGGTCGACCTGTACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3780
2641	TGGCTTCTATATTTTGTGATCTTTTGTGCTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT	2700	3721	CACGTGTGGAGCGGTCGACCTGTACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3780
2701	CCCTTATAGCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTT	2760	3781	AGAGCGGGGAGCAAAACGGGAGCGCTACTCTCCCGGAGACCTCTTTTCCACTTGAAGGG	3840
2701	CCCTTATAGCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTT	2760	3781	AGAGCGGGGAGCAAAACGGGAGCGCTACTCTCCCGGAGACCTCTTTTCCACTTGAAGGG	3840
2761	GGCCCAAGGAGTGTGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCGTGTCT	2820	3841	GTCTCAGGAGCGCGTGTGTATGCCAGGGGCGACGCTGTGGGAGTCTTCCGGGCGAGG	3900
2761	GGCCCAAGGAGTGTGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCGTGTCTCGGCGTGTCT	2820	3841	GTCTCAGGAGCGCGTGTGTATGCCAGGGGCGACGCTGTGGGAGTCTTCCGGGCGAGG	3900
2821	GGTAATGATCACTCTCTTTACTCTCACCCCGGGTATAAGACCTTTCTCAGCGGTTTTT	2880	3901	TGTGTGTCTCGGCGGCTGT	3960
2821	GGTAATGATCACTCTCTTTACTCTCACCCCGGGTATAAGACCTTTCTCAGCGGTTTTT	2880	3901	TGTGTGTCTCGGCGGCTGT	3960

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 12, 2004, 02:59:24 ; Search time 2189 Seconds
(without alignments)
11975.442 Million cell updates/sec

Title: US-09-980-559-1
Perfect score: 9711
Sequence: 1 accgccctaataggggcg.....ggctctcgagatcatgt 9711

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 2552756 seqs, 1349719017 residues
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	9711	100.0	9711	22	Nucleotide sequenc
2	9711	100.0	9711	22	Nucleotide sequenc
3	9707.8	100.0	9711	22	Infectious Hepatit
4	8970.4	92.4	9589	14	NANBH virus strain
5	8020.6	82.6	9678	24	Human HCV-related
6	6029.8	62.1	9511	14	NANBH virus strain
7	5800.8	59.7	9611	22	Nucleotide sequenc
8	5800.8	59.7	9611	22	Nucleotide sequenc

9	5692	58.6	9611	22	AAC86646	Nucleotide sequenc
10	5676	58.4	9611	22	AAC86648	Nucleotide sequenc
11	4625.6	47.6	9646	19	AAV59361	Hepatitis C virus
12	4625.6	47.6	9646	24	ABK87285	cDNA encoding hepa
13	4621.6	47.6	12980	24	ABK87286	Hepatitis C virus
14	4620	47.6	12980	19	AAV59364	Hepatitis C virus
15	4543.6	46.8	9622	25	AAAL54424	Hepatitis C virus
16	4534.4	46.7	9599	20	AAAX24832	Infectious hepatit
17	4523.2	46.6	9599	20	AAAX24832	Infectious hepatit
18	4523.2	46.6	9599	22	AAC86938	Nucleotide sequenc
19	4521.6	46.6	9599	22	AAAF23491	Infectious Hepatit
20	4506	46.4	9605	24	ABK91411	Hepatitis C virus
21	4506	46.4	9605	24	ABK91428	Hepatitis C virus
22	4506	46.4	9605	24	ABK91429	Hepatitis C virus
23	4506	46.4	9605	24	ABK91431	Hepatitis C virus
24	4506	46.4	9605	24	AAAD25332	Hepatitis C virus
25	4505.8	46.4	9595	20	AAAX24843	Infectious hepatit
26	4505.8	46.4	9595	22	AAC86939	Nucleotide sequenc
27	4504.4	46.4	9605	24	ABK91424	Hepatitis C virus
28	4504.4	46.4	9605	24	ABK91425	Hepatitis C virus
29	4504.4	46.4	9605	24	ABK91426	Hepatitis C virus
30	4504.4	46.4	9605	24	ABK91430	Hepatitis C virus
31	4504.4	46.4	9605	24	ABK91432	Hepatitis C virus
32	4504.4	46.4	9605	24	ABK91433	Hepatitis C virus
33	4504.2	46.4	9595	22	AAAF23492	Infectious Hepatit
34	4493	46.3	9608	24	ABK91427	Hepatitis C virus
35	4488.6	46.2	9609	24	AAD33038	HCV-S1 full-length
36	4486.2	46.2	9436	15	AAQ63499	Blood transmissibl
37	4468.8	46.0	9502	15	AAQ74770	Hepatitis C virus
38	4452.8	45.9	9487	16	AAQ79140	Hepatitis C virus
39	4449.2	45.8	9413	24	AAD25517	Hepatitis C virus
40	4449.2	45.8	9413	25	AAD49655	Hepatitis C virus
41	4447.6	45.8	9413	25	AAAL53723	Hepatitis C virus
42	4446	45.8	9413	16	AAAT03960	Partial HCV non-st
43	4444.4	45.8	9413	16	AAQ81559	Hepatitis C virus
44	4436.4	45.7	9413	16	AAQ80498	DNA encoding HCV p
45	4433.8	45.7	10803	25	ABX10617	MKO-Z nucleotide s

ALIGNMENTS

RESULT 1
AAC86644
ID AAC86644 standard; DNA; 9711 BP.
AC AAC86644;
XX
XX
XX 02-APR-2001 (first entry)
XX
XX
XX Nucleotide sequence of infectious Hepatitis C virus strain HC-J6CH.
XX
XX HCV; HCV strain HC-J6CH; HCV genotype 2a; antiviral; vaccine; ss.
XX
XX Hepatitis C virus.
XX
XX Key Location/Qualifiers
XX CDS 341..942
XX /*tag= a
XX /*product= "Hepatitis C virus protein"
XX
XX WO200075338-A2.
XX
XX 14-DEC-2000.
XX
XX 02-JUN-2000; 2000WO-US15446.
XX
XX 04-JUN-1999; 99US-0137693.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Yanagi M, Bukh J, Emerson SU, Purcell RH;
XX

8941	GATGACACACTTTCTTCTCCATTCTCATGGCTCAAGACAGCGCTGGACCAAGACCTCAACTT	9000
8941		
8941	GATGACACACTTTTACTCTCAATCTCATGGCCCAAGACACCCCTGGACCAAGACCTTAACTT	9000
9001	TGAGATGACGAGCGGTGTACTCGTGAGTCCCTTGGACCTCCCGAGCTATTAATTGAAG	9060
9001		
9001	TGAGATGACGAGCGGTGTACTCGTGAGTCCCTTGGACCTCCCGAGCTATTAATTGAAG	9060
9061	GTTACATGGGCTTGACCGCTTTTCTCTGCACACATACACTCCCCACGAACCTGACCGGT	9120
9061		
9061	GTTACACGGCTTGACCGCTTTTCACTGCAACATATACTCCCCACGAACCTGACCGGGT	9120
9121	GGCTTCAGCCCTCAGAAAACCTTGGGGCGCCACCCCTCAGAGGTGGAAAGACCGGGCACG	9180
9121	GGCTGCAGCCCTCAGAAAACCTTGGGGCACCAACCCCTCAGAGCGTGGAAAGCTCGGCGCG	9180
9181	TGCAGTCAGGGGTCCCTCATCTCCCGTGGGGGAGAGCGGCGTTTCGGGTGCATATCT	9240
9181		
9181	TGCGGTTAGGGCGTCCCTCATCTCCCGTGGAGGAGAGCGGCGCTCTCGGTCGGTATCT	9240
9241	CTTCAAATTGGGCGGTGAAGACCAAGCTCAAACTCACTCTCATTTGCCGGAAGCGCGCTCCT	9300
9241		
9241	CTTCAACTGGGCGGTGAAGACCAAGCTCAAACTCACTCTCATTTGCCGAGGCACGCCACT	9300
9301	GGATTTATCCAGCTGGTTTCAACGTCGGCGCCGGGGGGGCGACATTTATCAGCGGTGTC	9360
9301	GGATTTATCCAGTTGGTTTCAACGTCGGCGCCGGGGGGGCGACATTTATCAGCGGTGTT	9360
9361	CGGTCCCGACCCCGCTTATTCCTTTGGCTACTCCTACTCTTTTGTAGGGGTAGGCT	9420
9361		
9361	CGGGCCCGACCCCGCTTATTTACTCTTTGGCTACTCTCTCTTTTCGTAGGGGTAGGCT	9420
9421	TTTCTACTCCCGCTCGGTAGAGCGGCACACATTTAGCTACACTCCATAGCTAACTGTCC	9480
9421		
9421	CTTCTACTCCCGCTCGGTAGAGCGGCACACATAGCTACACTCCATAGCTAACTG---	9477
9481	CTTT	9540
9478	-----TT	9492
9541	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCTTTTCTCTTTTCTCTTTTCTTTACCTATTT	9600
9493		
9493	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCTTTTCTCTTTCTCTTCTACTCTCTT	9552
9601	TACTTTCTTCTCGTGCGCTCCATCTTAGCCCTAGTCAACGGCTAGCTGTGAAGGTCCGT	9660
9553	TACT---TTTTCTGGTGGCTCCACTTAGCCCTAGTCAACGGCTAGCTGTGAAGGTCCGT	9609
9661	GAGCGGCATGACTGCAGAGAGTGCCTTAACCTGGTCTCTCTCGAGATCATGT	9711
9610	GAGCGCATGACTGCAGAGAGTGCCTTAACCTGGTCTCTCTCGAGATCATGT	9660

Copied from 09540843 on 05/19/2004

Db	5641	GCACATGTGGAACTTCATTAGCGCATCCAAATACCTCGCGGGAGTGTCAACGCTGCGAG	5700		Db	6721	CTGGGTGGACGGAGTGCAAAATCCATAGGTTTGTCTCCCAAGCAAAAGCCGTTTTTCCGGGA	6780
Qy	5701	GAACCTGAGTAGTTCATGATGCGGTTCAGTGGCGCCTCACAGTCCGTGTCAAC	5760		Qy	6781	TGAGGTCTCGTTCTCGTTGGGCTTAATTCATTGTCGTGGGTCCAGCTTCCTTTCGGA	6840
Db	5701	AAACCCCGCTGTGGCTTCCATGATGCGATTCAGTGGCGCCTCACAGCCGTTGTCAAC	5760		Db	6781	TGAGGTCTCGTTCTCGTTGGGCTTAATTCATTGTCGTGGGTCTCAGCTTCCTTGTGA	6840
Qy	5761	AAGCACCATATCTCTCTCAACATTTTGGGGGCTGGCTAGCATCCCAATTCGACAC	5820		Qy	6841	CCTGAAACCCGACACAGACGTATTGATGTCATGCTAAACAGATCCATATCATACGGC	6900
Db	5761	TAGCACCATATCTCTCTCAACATTTTGGGGGCTGGCTAGCATCCCAATTCGACAC	5820		Db	6841	CCTGAAACCCGACACAGACGTATTGATGTCATGCTAAACAGATCCATATCATACGGC	6900
Qy	5821	CGCGGGGCACTGCGGCTTCGTTGATGTCGCTAGTGGAGAGTGGCGGTAGGCAAGTAGG	5880		Qy	6901	GGAGGTGAGCGCGCGGCTTTAGCGGGGGTACCCCATCCGAGGCAAGCTCTCTCGC	6960
Db	5821	CGCGGGGCACTGCGGCTTCGTTGATGTCGCTAGTGGAGAGTGGCGGTAGGCAAGTAGG	5880		Db	6901	GGAGGTGAGCGCGCGGCTTTAGCGGGGGTACCCCATCCGAGGCAAGCTCTCTCGC	6960
Qy	5881	CTTAGGTAGGTGCTAGTGACATCTTGGGAGGATGTTGGGCGCATTTCCGGGGCTCT	5940		Qy	6961	GAGCCAGCTATCGGCAACCATCGCTGCGAGCCCTGCGACACCCACGGAAGACCTATGA	7020
Db	5881	CTTAGGTAGGTGCTAGTGACATCTTGGGAGGATGTTGGGCGCATTTCCGGGGCTCT	5940		Db	6961	GAGCCAGCTATCGGCAACCATCGCTGCGAGCCCTGCGACACCCACGGAAGACCTATGA	7020
Qy	5941	CGTCGCAATCAAGATCATGTCTGGGAGAGCCCTCCATGGAGGATGTCGTCAACTTGCT	6000		Qy	7021	TGTGGACATGTTGGATGCTTAAACCTGTTTATGGGGGGCGATGTGAATCGGATAGAGTCTGG	7080
Db	5941	CGTCGCAATCAAGATCATGTCTGGGAGAGCCCTCCATGGAGGATGTCGTCAACTTGCT	6000		Db	7021	TGTGGACATGTTGGATGCTTAAACCTGTTTATGGGGGGCGATGTGAATCGGATAGAGTCTGG	7080
Qy	6001	GCTTGGAAATCTGTCTCGGGTGTCTTGGTAGTGGAGTCAATCTGGCGGCGCATTTCTGGG	6060		Qy	7081	GTCCAAAGTGTCTGTCGAGCTCTCTCGACCCCAATGCTCGAAGAAAGGAGCGACCTTGA	7140
Db	6001	GCTTGGAAATCTGTCTCGGGTGTCTTGGTAGTGGAGTCAATCTGGCGGCGCATTTCTGGG	6060		Db	7081	GTCCAAAGTGTCTGTCGAGCTCTCTCGACCCCAATGCTCGAAGAAAGGAGCGACCTTGA	7140
Qy	6061	CCGACAGTGGGAGCGGGGAGGCGGCTCCATGATGATGATGATGATGATGATGATGATG	6120		Qy	7141	GCCTTCGATACCATCAGAAATACATGCTCCCAAGAAAGAGGTTCCCAACAGCTTTACCGGC	7200
Db	6061	CCGACAGTGGGAGCGGGGAGGCGGCTCCATGATGATGATGATGATGATGATGATGATG	6120		Db	7141	GCCTTCGATACCATCAGAAATACATGCTCCCAAGAAAGAGGTTCCCAACAGCTTTACCGGC	7200
Qy	6121	TTCCAGAGGAAATCACGCTCGCCGCCACCTACTAGTACGAGGTCGATGCTCGAGCG	6180		Qy	7201	CTGGGCAACGCTGATTAACAACCCGCTTGTGGAATCGTGGAAAGGCGAGTTACCA	7260
Db	6121	TTCCAGAGGAAATCACGCTCGCCGCCACCTACTAGTACGAGGTCGATGCTCGAGCG	6180		Db	7201	CTGGGCAACGCTGATTAACAACCCGCTTGTGGAATCGTGGAAAGGCGAGTTACCA	7260
Qy	6181	TGTGACCCAGTGTCTGCTTGGCTTACATTAACATAGCTGCTCAGAAAGCTCCTCAATTTG	6240		Qy	7261	ACGGGCACTGTGCGGGTGTCTCTCTCTAGGAAACCCCGGAGCTCCCGC	7320
Db	6181	TGTGACCCAGTGTCTGCTTGGCTTACATTAACATAGCTGCTCAGAAAGCTCCTCAATTTG	6240		Db	7261	ACGGGCACTGTGCGGGTGTCTCTCTCTAGGAAACCCCGGAGCTCCCGC	7320
Qy	6241	GATTACTGAGGACTGCGGCTTCCATGCGGGGCTCTGCTGCTGCTGCTGCTGCTGCTG	6300		Qy	7321	AAGGAGCGCGGACAGTGGGCTTAAGTGAAGACTCTCATAGAGAGTGCCTTCAACAGCT	7380
Db	6241	GATTACTGAGGACTGCGGCTTCCATGCGGGGCTCTGCTGCTGCTGCTGCTGCTGCTG	6300		Db	7321	AAGGAGCGCGGACAGTGGGCTTAAGTGAAGACTCTCATAGAGAGTGCCTTCAACAGCT	7380
Qy	6301	GGTTTGACCAATCTTAACAGACTTTAAATAATGCTGACCTTCCAAATTTATTTCCAAAGAT	6360		Qy	7381	GGCCATTAAGTCTTTGGGCGAGCCCCCGGAGCGGAGTTAGGCTTTTCCAGCGGGGC	7440
Db	6301	GGTTTGACCAATCTTAACAGACTTTAAATAATGCTGACCTTCCAAATTTATTTCCAAAGAT	6360		Db	7381	GGCCATTAAGTCTTTGGGCGAGCCCCCGGAGCGGAGTTAGGCTTTTCCAGCGGGGC	7440
Qy	6361	GCCCGGCTTCCCTTGTCTGTCGTAAGAGGGGTAACAAGGCGTGTGGGCGCGGCACTGG	6420		Qy	7441	GGGCGTTCGCGATTCGCGGAGTCCAGAGCTCTCTGATGCTGCTGCTGCTGCTGCTGCTG	7500
Db	6361	GCCCGGCTTCCCTTGTCTGTCGTAAGAGGGGTAACAAGGCGTGTGGGCGCGGCACTGG	6420		Db	7441	GGGCGTTCGCGATTCGCGGAGTCCAGAGCTCTCTGATGCTGCTGCTGCTGCTGCTGCTG	7500
Qy	6421	CATCATGACACAGGCTGCTTGGCGGCGCAATATCTTGGCAATGTCGCTTGGGCTC	6480		Qy	7501	TTCCATCTCTTCCATGCTTCCGAGGGGAGCTTGGAGATCCAGAGCTGAGAGCTGGA	7560
Db	6421	CATCATGACACAGGCTGCTTGGCGGCGCAATATCTTGGCAATGTCGCTTGGGCTC	6480		Db	7501	TTCCATCTCTTCCATGCTTCCGAGGGGAGCTTGGAGATCCAGAGCTGAGAGCTGGA	7560
Qy	6481	CATGAGATCAAGGCGCTAAGACTGCTGATGATATCTGGGAGGAGCTTTTCCATCAA	6540		Qy	7561	GAAGGTAGAGGCGGCAACCCCGGAGGGGGTGGAGCTCCCGGCTCGGAGCTCGG	7620
Db	6481	CATGAGATCAAGGCGCTAAGACTGCTGATGATATCTGGGAGGAGCTTTTCCATCAA	6540		Db	7561	GAAGGTAGAGGCGGCAACCCCGGAGGGGGTGGAGCTCCCGGCTCGGAGCTCGG	7620
Qy	6541	TTGTTACACGAGGGGCGAGTGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	6600		Qy	7621	GTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	7680
Db	6541	TTGTTACACGAGGGGCGAGTGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	6600		Db	7621	GTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	7680
Qy	6601	GAGGTGCGGCTTCCAGAGTACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	6660		Qy	7681	CTGAGCGGGGCTTAATAACTCTCTTGTAGTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	7740
Db	6601	GAGGTGCGGCTTCCAGAGTACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	6660		Db	7681	CTGAGCGGGGCTTAATAACTCTCTTGTAGTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	7740
Qy	6661	AGGACTCAACATGATTAAGTCCCGTGGCAATCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	6720		Qy	7741	CTTGAGCAACTCCTGTTGCGATATCAACAAGGTGTAATGTAACCAACAAAGAGCGC	7800
Db	6661	AGGACTCAACATGATTAAGTCCCGTGGCAATCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	6720		Db	7741	CTTGAGCAACTCCTGTTGCGATATCAACAAGGTGTAATGTAACCAACAAAGAGCGC	7800
Qy	6721	CTGGGTGAGCGAGTGCAGATCCATAGGTTTGGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	6780		Qy	7801	CTCAGTAAAGGCTTAAAGAGTAACTTTGATAGGATCCAGAGTCTGCTGCTGCTGCTGCTGCTG	7860
Db	6721	CTGGGTGAGCGAGTGCAGATCCATAGGTTTGGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	6780		Db	7801	CTCAGTAAAGGCTTAAAGAGTAACTTTGATAGGATCCAGAGTCTGCTGCTGCTGCTGCTGCTG	7860

QY	3481	AGTGGTGAAGCATGACCGGGCGCGACAGACAGAGGCTGGGAAAATTTCAGGTCTCTGTC	3540	4561	CGCGCGCGCCCTTCGGGGTATCGGCTTGAACTCAGTGGCATACTACAGAGGGTGGACGT	4620
DB	3481	AGTGGTGAAGCATGACCGGGCGCGACAGACAGAGGCTGGGAAAATTTCAGGTCTCTGTC	3540	4561	CGCGCGCGCCCTTCGGGGTATCGGCTTGAACTCAGTGGCATACTACAGAGGGTGGACGT	4620
QY	3541	CACAGTCACTCAGTCTCTTCGGAACATCATCTCTCGGGGTTTGTGACATCTCTACCA	3600	4621	CTCCGTATATCAAACTCAGGAGAGCTAGTGTCTGCCACCGACCGCCCTCATGACAG	4680
DB	3541	CACGGTCACTCAGTCTCTTCGGAACATCATCTCTCGGGGTTTGTGACATCTCTACCA	3600	4621	CTCCGTATATCAAACTCAGGAGAGCTAGTGTCTGCCACCGACCGCCCTCATGACAG	4680
QY	3601	TGAGCTGGCAACAGAGCTCTGCGCGCTCACGGGGTCCGGTCCAGAGATGACTCCAG	3660	4681	GTATATCGGGGACTTTGACTCTCGTCTGATCGACTGCAAGTAGGGTCTCAAGTCTGTA	4740
DB	3601	TGAGCTGGCAACAGAGCTCTGCGCGCTCACGGGGTCCGGTCCAGAGATGACTCCAG	3660	4681	GTATATCGGGGACTTTGACTCTCGTCTGATCGACTGCAAGTAGGGTCTCAAGTCTGTA	4740
QY	3661	TGCTGAGGGGGAATAGTAGGGTGGCCAGCCCTCTGGGACTAAATCTTTGAGCGGTG	3720	4741	CTTCAGTTTATAGACCCACATTCACATTAACACACAGATGTGCTCCCTCAAGACGCTCTC	4800
DB	3661	TGCTGAGGGGGAATAGTAGGGTGGCCAGCCCTCTGGGACTAAATCTTTGAGCGGTG	3720	4741	CTTCAGTTTATAGACCCACATTCACATTAACACACAGATGTGCTCCCTCAAGACGCTCTC	4800
QY	3721	CACGTGTGAGAGCGGTGACCTGTACCTGTGTCACGCGGAAAGCTGTATCTATCCCGCTCG	3780	4801	ACGTAGCCAGCCCGGGTTCGCAACGGGTAGGGGAGAGCTGGGCAATTTATAGGTATGTTTC	4860
DB	3721	CACGTGTGAGAGCGGTGACCTGTACCTGTGTCACGCGGAAAGCTGTATCTATCCCGCTCG	3780	4801	ACGTAGCCAGCCCGGGTTCGCAACGGGTAGGGGAGAGCTGGGCAATTTATAGGTATGTTTC	4860
QY	3781	AAGACGGGGGACAAAGCGGAGCGGTACTCTCTCCCGAGACCTCTTCCACCTTGAAGG	3840	4861	CACTGTGTGAGCGAGCTTACAGGCTCAGGAACTGTTTCAAGTGTAGTGTCTGTAGTCTACACGC	4920
DB	3781	AAGACGGGGGACAAAGCGGAGCGGTACTCTCTCCCGAGACCTCTTCCACCTTGAAGG	3840	4861	CACTGTGTGAGCGAGCTTACAGGCTCAGGAACTGTTTCAAGTGTAGTGTCTGTAGTCTACACGC	4920
QY	3841	GTCTCTGAGGAGCGGTGCTATGCGCCAGGGGCGACGCTGTCCGGAGTCTTCGGGCGAGC	3900	4921	AGGGGCGCGCATGTATGAGCTCACACCATCGGAGACACCGCTCAGGCTCAGGCGGTATTTT	4980
DB	3841	GTCTCTGAGGAGCGGTGCTATGCGCCAGGGGCGACGCTGTCCGGAGTCTTCGGGCGAGC	3900	4921	AGGGGCGCGCATGTATGAGCTCACACCATCGGAGACACCGCTCAGGCTCAGGCGGTATTTT	4980
QY	3901	TGTGTGCTCTCGGGGCTGGCTAAGTCCATAGATTTTCATCCCGTTGAGACATCGACAT	3960	4981	CAACACGCGCGTTCCTCTGTGTCGCAAGACCATCTTGAGTTTGGGAGGCAATTTT	5040
DB	3901	TGTGTGCTCTCGGGGCTGGCTAAGTCCATAGATTTTCATCCCGTTGAGACATCGACAT	3960	4981	CAACACGCGCGTTCCTCTGTGTCGCAAGACCATCTTGAGTTTGGGAGGCAATTTT	5040
QY	3961	CGTCAACGGGTCCTCCACCTTTAGTGAACACAGACACACCTGTCTGCGCCAGACCTA	4020	5041	CGGCTCTCACACATAGATGCCCCACTTCTTCCCAACAAAGACATCGGGGAAAATTT	5100
DB	3961	CGTCAACGGGTCCTCCACCTTTAGTGAACACAGACACACCTGTCTGCGCCAGACCTA	4020	5041	CGGCTCTCACACATAGATGCCCCACTTCTTCCCAACAAAGACATCGGGGAAAATTT	5100
QY	4021	TCAGGTGCGGTATTTGATGCCCCGAGTGGCACTGCAAGACCAACAGTCCCTGTGCG	4080	5101	CGCATACTTAAACAGCGCTACAGGCTACAGTGTGCGTACAGGCTTCCCTCTCTCGCTC	5160
DB	4021	TCAGGTGCGGTATTTGATGCCCCGAGTGGCACTGCAAGACCAACAGTCCCTGTGCG	4080	5101	CGCATACTTAAACAGCGCTACAGGCTACAGTGTGCGTACAGGCTTCCCTCTCTCGCTC	5160
QY	4081	ATATGTCTCAGGGGTATAAGTGTAGTGTGTTAATCTCAGTGGCTGCCACCTGGG	4140	5161	CTGGAGCGTCTATGTGGAGTGTTCAGTCTCAAGCGCACTCGTGGGGCCCAACACC	5220
DB	4081	ATATGTCTCAGGGGTATAAGTGTAGTGTGTTAATCTCAGTGGCTGCCACCTGGG	4140	5161	CTGGAGCGTCTATGTGGAGTGTTCAGTCTCAAGCGCACTCGTGGGGCCCAACACC	5220
QY	4141	GTTTGGGGGTACTTGTCTAAGGCAATGGCTCAATCCCAATTTAGGACTGGAGTCA	4200	5221	TCCTCTGTACCGTTTGGGCTCTGTACCAACAAAGTCACTCCCTCACACATCTCTGTGACGAA	5280
DB	4141	GTTTGGGGGTACTTGTCTAAGGCAATGGCTCAATCCCAATTTAGGACTGGAGTCA	4200	5221	TCCTCTGTACCGTTTGGGCTCTGTACCAACAAAGTCACTCCCTCACACATCTCTGTGACGAA	5280
QY	4201	GACTGTGACGACCGGGCGCCATCACGTACTCCACATATGGCAATTTCTCGCGATGG	4260	5281	ATACATCGCCACCTCATGCAAGCGACCTTGGAGTCTATGACGACATGGGTCTTGGC	5340
DB	4201	GACTGTGACGACCGGGCGCCATCACGTACTCCACATATGGCAATTTCTCGCGATGG	4260	5281	ATACATCGCCACCTCATGCAAGCGACCTTGGAGTCTATGACGACATGGGTCTTGGC	5340
QY	4261	GGGCTGCGGAGCGCGCTATGACATCATATGCGATGAATGCGCTGGACGC	4320	5341	AGGGGAGTCTTGGCGCGCTCGCGGTATTCCTTGGCGACCGGGGTGTGTTGATCAT	5400
DB	4261	GGGCTGCGGAGCGCGCTATGACATCATATGCGATGAATGCGCTGGACGC	4320	5341	AGGGGAGTCTTGGCGCGCTCGCGGTATTCCTTGGCGACCGGGGTGTGTTGATCAT	5400
QY	4321	TACCATCTCTGGGATCGGAACAGTCTCTGATCAAGCAGAGACAGCTGGGGTCAAGT	4380	5401	CGGCGCTTGCACATTAACACGAGCGCTGTGCGCGGACCAAGAGGTCTCTATGCA	5460
DB	4321	TACCATCTCTGGGATCGGAACAGTCTCTGATCAAGCAGAGACAGCTGGGGTCAAGT	4380	5401	CGGCGCTTGCACATTAACACGAGCGCTGTGCGCGGACCAAGAGGTCTCTATGCA	5460
QY	4381	AACGTGTGCTGCTACAGCTACGCGCTGCGTCAAGTGAACACCCCGCCACCAATAGA	4440	5461	GGCTTTTGTATGAGTGGAGGATGTCCTTAGGGCGGCTCTCATTTGAAGAGGGGCGCG	5520
DB	4381	AACGTGTGCTGCTACAGCTACGCGCTGCGTCAAGTGAACACCCCGCCACCAATAGA	4440	5461	GGCTTTTGTATGAGTGGAGGATGTCCTTAGGGCGGCTCTCATTTGAAGAGGGGCGCG	5520
QY	4441	GGAGGTGGCCCTTGGCAGGAGGCGAGATCCCTTTCTATGGAGGGCGATTTCCCTGTG	4500	5521	GATAGCCGAGATGCTGAAGTCAAGATCAAGGCTTTATGACGAGGCTTCAAAACAGC	5580
DB	4441	GGAGGTGGCCCTTGGCAGGAGGCGAGATCCCTTTCTATGGAGGGCGATTTCCCTGTG	4500	5521	GATAGCCGAGATGCTGAAGTCAAGATCAAGGCTTTATGACGAGGCTTCAAAACAGC	5580
QY	4501	TTATCATCAAGGGAGGAGACATCTGATCTTCTGCCATTTCAAGAAAAGTGTGACGAGT	4560	5581	TCAGACATACAAACCACTGTGCGGCTTCTAGGCGGAGGTAGAACATTTCTGGGCGAA	5640
DB	4501	TTATCATCAAGGGAGGAGACATCTGATCTTCTGCCATTTCAAGAAAAGTGTGACGAGT	4560	5581	TCAGACATACAAACCACTGTGCGGCTTCTAGGCGGAGGTAGAACATTTCTGGGCGAA	5640
				5641	ACACATGTGGAATCTTCAATAGCGGATCCAATATCTCGCAGGACTATCAACACTGCCAGG	5700

3341	TCTTTTGTGCACTCTACCA	CGAATGGGCCATTTTGGCCCTGCTCCTACTCGACCTGCC	2400	
3401	CGCCTTTTCGA	CTGTCTTTCTCCACCTCCACAAACAATCGTGGAGCTACAATTCATGTA	2460	
3401	CGCTTTATCGA	CTGGCTTCTCCACCTCCACAAACAATCGTGGATGATCAATACATGTA	2460	
3461	TGGCCTTATCA	CTCGCCCTCACAATAACATCGTCCGATGGAGTGGGTAAATCACTCTTAT	2520	
3461	TGGTCTGT	CACCTGCGCCTCACAATAATACATCGTCCGATGGAGTGGGTAGTACTCTTAT	2520	
3521	CTGTCTTTAG	CGGACGCGAGGTTTGGCCCTGCTTATGGATGCTCATCTTGTGGGCCA	2580	
3521	CTGTCTTTAG	CGGAGCTAGGGTCTGTGCTGCTTATGGATGTTTATCTTGTATGGGCCA	2580	
3581	GGCCGAAGCAGCA	CTAGAGAACTGGTCACTTGCA	CGCTCGGAGCGCAGCTAGCTGCAAA	2640
3581	GGCCGAAGCAGCA	CTAGAGAACTAGTGGTCTTGCA	CGCTCGGAGCGCGCTAGCTGCAAA	2640
3641	TGGCTTCTTA	TATTTTGTGTCATCTTTTTCGTGGCTGCTTGGTGTACATCAAGGGTCGGGTACT	2700	
3641	CGGCTTCTTA	TATCTTTTGTGTCATCTTTTTCGTGGCTGCTTGGTGTACATCAAGGGTCGGGCAGT	2700	
3701	CCCTTTAGCT	TACTTATTCCTTCAC	TGGCTGTGCTTTCCTTTAGCTACTCTGCTCTAGCATT	2760
3701	CCCTTTGGCT	GTATTCCTCACCGGCTATGGTCTTTCTGGCTATTCCTCTAGTACT	2760	
3761	GCCCAACAGGCT	TATGCTTATGACGATCTGTGCAT	TGGCCAGATAGGAGCGCTCTGCT	2820
3761	GCCCCAGCAGGCT	TATGCTATGAGGCATCTGTGCA	CGGACAGATAGGCGCGCTTTGTT	2820
3821	GGTAATGATCA	CTCTCTTTACTCTCACCCCGGGGTATAAGACCCCTTCAGCCGGTTTT	2880	
3821	GGTACTAA	TATCCCTTTTACACTCACCCCGGGGTATAAGACCCCTTCAGTGGGTGCT	2880	
3881	GTGTGGTGTG	CTATCTTCTGACCTCGGGGAAAGCTATGGTCTCAGGAGTGGGCACCAAC	2940	
3881	GTGTGGTGTATG	CTATCTTCTGACTCTGGGGGAGGCCATGGTCCAGGAGTGGGCACCATC	2940	
3941	TATGACAGGT	CGCGGTGGCCGTATGGCATCATATGGGCCGTGCCCATATTCACCCAGG	3000	
3941	CATGACGGCG	CGCGGTGGCCGATGGCATCATCTGGGCCGCCACATATTCCTGCCCGGG	3000	
3001	TGTGGTGT	TGACATTAACCAAGTGGCTCTTGGCGGTGCTTTGGGCGCTTCTACCTCCATAA	3060	
3001	CGTGGTGT	TGACATTAACCAAGTGGCTTTGGCGCGCTTGGGTCTGGCTATCTCTTAAG	3060	
3061	AGGTCTTTG	AGCGGTGCGCTGCTGCTCGT	CAGGCTCACGCTCTATGAGGATGTGCAC	3120
3061	AGGTCTTTG	AGCGGTGCGCTGCTGCTCAGGCTCACGCTCTATGAGGATGTGCAC	3120	
3121	CATGGCAAGGCA	TCTCGCGGGGGCAGGTACGTC	CAGATGCGCTACTAGCCCTTGGCAG	3180
3121	TATAGCGAAGCA	CCTCGCAGGGGGTAAGTACGTC	CAAAATGCGCTATTAGCCCTTGGTAG	3180
3181	GTGACGTGGCA	CTTATCTATGACGACCTCACCCCTATGT	CGAATGGGCTCTCTAGTGG	3240
3181	GTGACGTGGCA	CTTATCTATGATCAGCTCATCCCTCATGT	CGAATGGGCTGCCAGTGG	3240
3241	CCTGCGGGA	CTTGGCGCTTATCATCTT	CAGTCCGATGGAAGAAGT	3300
3241	CCTGCGGGA	CTTGGCGCTTATCATCTT	CAGTCCGATGGAAGAAGT	3300
3301	CATTGTCTGGG	GACGAGACAGCTGCTTGTGGGACATTTTA	CACCGGACTTCCCGTGC	3360
3301	TATCGTCTGG	GAGCGGAGACGCTGCA	TGCGGGGACATCTTACACGGACTTCTCTGTGC	3360
3361	CGCCCGA	CTTGGTCTGGGAGGTCTCTCTTGGGCCAGCTGATGGCTTACCTCCAGGGGTG	3420	
3361	CGCCGRC	CTCGTCTGGGAGATCTCTCTTGGGCCAGCTGATGGCTTACCTCCAGGGGTG	3420	
3421	GAGTCTTCT	CGCCCCCATCACTGCTTACGCCCAGCAGACAGTGGCCCTTTTGGGCACCAT	3480	
3421	GAGGCTTCT	TGGCCCCCATCACCGCTTACGCCCAGCAGACAGAGTCTTTGGTGTAT	3480	

```

/db_xref="GI:6707280"
/translation="MSTNPKPQKTKRNTNRRPQDKVFGGQIVGGVLLPRLRRGPR
GVRATKTSRSPRRRQIPKDRSTGRSGKPGYPMPLVNEGLGWAGMLSPRG
SRPSCGPNDRSRNVKVIDLTCGFADLAMYIPVVGAPLGGVATALAHGVRLED
GINPATGNLPGFSFSLALSLSCITVPVSAQKNTSSGTYMTDNCANDSITWQLQA
VLVHPGCVFCXVXXSRWIPVSNVAVRQKALTQGLRLTHDMVMSATLCSALY
VGLDCGVMLAAQITVSEPHFWOECNSIYPGAIITGORMAWMMNWSPTATMIL
AVAMRVEPIIDIIISAGHVMFGLAYFSMOGAKVAVVILLAAQVADANTXXGGA
ARNAYVSLNFKXGAKONQILINTNGSMHINTALNCNDSLANTGLFLAXLYXNRNNS
CGPRLSVLSEAFRIHGTLOVEDNVTNPMRPHYCHYPPKQCSIVPASPVGVP
YCFPSFVVGTDRNGVETTYWGENETDFLLNRPPOGSGWFGCTNMGSTFKTC
GAPCTRDFNASDLCCTCFKRPHEATYIKGSGFWLTPRCLVDYFRLWHYPC
TNYNITFKRMVYGVYHRLTAACNFRGDRCDLDRSOLSLPSTHTEWAILPCS
YSDLPALSTGLHLHONIVDVVOMYGLSPALTKYIVRWEMVLLPILLADARVCAQLW
MFLMQAARALEKLVLAHAASAACNGFLYFVFEVAAYIKRVAAPLAAYSLGLW
SFCLLLVLPQQAAYEASVHGIGNALLVLITLFTLTPCYKTLISRLWLWLYLLTL
GEAWQEPASQWARGROGIIWAATIFCGVVFVDITKMLLAALSGYLLRGLALTRVP
YFVRAHALLMRKITAGLGGKYQWALLALGRTWTYIYDHLTPMSDWAASGLRLLA
VAVEPIIFSPMEKKVIVMGAETAACDGLHGLFVPSARLGREILLGPADGYTSGWRL
APITAVAOOTRGLLGAIVVSMTRDKTEQAGBIQVLSVTQSLGTSISGLVWTVHG
AGNKTLAGSRGPVTOYSSAEGDLGWPSPPGCTKLEPCTCGAVDLVLRNADVIPA
RRGDKRGALLSPRLSTLKSISGGPVLCPRGHANGVIFRAACVSRGVAKSIDFIVET
LDIVTSRPTSDNTPPAVPQIYQVGLHAPTSKSTKVPVAYASQGVKVLVLPVS
AATLPGFAYLSKAHGINPRTGVRTVITTEATYTSYKFLADGCGAGGAYDIIICD
ECHANDAITLIGITVLDQAEATAGVRLTVLATATPSTVTPHPNIEEVALHGEPI
FYKALPLSVIKGGRHLIFCHSKKCKDELGAALRGMLNAVAYYRLDVSAPAOQDV
VVATDALMTGTDSDSVIDCNVAVTOVVDLSLOPTFTITVTPQDAVSRQGR
TGRGLGIYRYSTGERASGMFDSVILCEYDAGAAVYELTAEITVRLRYNPTPL
PQCDELFEWAVFTGLTHIDHFLUSQTKSGENFAYLVAQATVCARAKAPPBNDV
MKCLRLTLTLVPTPPLLYRLSGVNEVTLTHPVTKYIATCMQADLEVMSTVTLAG
GVLAAYVCLATGVQSIIGRIHQINQAVAPDKEVLEAFDEMEBCASRAALIIEGQ
RTAELMKXIOGLQOAKOQADIOAPVQOASPKVQEPKAWHMNFIISGIOYLAGLST
LPGNPAVASAFASLSTPLSTLILNLLGMLASQIAPAGATGPVVSGLVGA
VGSVGLGVLDILAGYAGISGALVAFKVMGSRKPSMEDVNLPLGILSPALVGV
ICAAILRRHVGEQAVQMNKLIIPASRGNHVAPTHYVTSDESQORVTLGLSITIT
SLRLRHMNITDEPCIPACAGSLNDRVMDVCTITLDFKNWLTSLKFPKMPGLFISQ
KGYRWAGTGTMTTRCSGASISGNVRLGSMRTIGPRTCMNTWQGTPTINCYTEGQ
VQHPAPKPKPFRDEVSFCVLSNVGSQJLPDEPDTDLMSLMDPDSHITAAA
ARLARGSPSASASASQSPASRATCKGTVDVMDANLFGMDVTRISSES
KVVLDLSDLPMAERSDLEPSIPSEMLPNNRFPALPAWARPNNPLVESWKRPDY
QPTVAGCALPPKAPPPRRRTVGLSESTIGDALQQLAKITFGPPPPSGDGLS
TGADAAAGCGQASPELALSETSTSMPELEGPDPDLERVELQALPOGSEVAP
GSDSGWSCTSEDDSVCCSSYSMTGALIIPSGPEEBKLIPNLSNLLRYHNKV
CTTSRTASLRAKVFTRQVLDAYVDSPLKDIKLAASRVSAARLITLSEACOLTPHS
ASRVGFGAKVRSLSGRAVNHKSVWKDLLEDSTQPIPTTMAKNFVCEVPTGGK
KAARLVYDPLSVRVERKMAVYDQKUPQAVMGASYGQSPAPQORVBFLLKAWAEK
DPMGFSYDFRDSVTTERDIRTEESIVQACSLPEEARTAIHSLETERLYVGGPMLNSQ
GOACGYRRCRASGLVLTSMGNTITCYIKALAAKAAAGIVAPTMLVCGDDLVVISESQ
TEEDERNLRAFTEAMTRYSAPPDPPEYDLELITSCSSNVSVVALGQRRRYLYTR
DPTTPXARAWEITVRHSPVNSWGLNIOYAPTIWVRVLMWTHFYSLMAQDTLDONL
FEMYGANVSVSLDLPALIERLHGLDASLHLYTHPHELTRVAALRKLGAAPLRWKS
HARVASLISRGGRAAVCGRYLFWNAVKLKLTPLEARHLIDLSMFTVAGGGDI
YHSVLRARLRLGLLGLLVGVGLFLPAR"
BASE COUNT 1963 a 2698 c 2698 g 2131 t 44 others
ORIGIN

```

```

Query Match 85.8%; Score 8329.4; DB 14; Length 9661;
Best Local Similarity 91.1%; Pred. No. 0;
Matches 8842; Conservative 41; Mismatches 777; Indels 51; Gaps 2;

QY 1 ACCCGCCCTAATAGGGCGGACACTCCGCGCATGAATCACTCCCTGTGTAGGAACACTGT 60
DB 1 ACCCGCCCTAATAGGGCGGACACTCCGCGCATGAATCACTCCCTGTGTAGGAACACTGT 60

QY 61 CTTACGCGAAGCGCTAGCCATGGCGGTAGTAGTGTGCTACAGCCTCCAGGCC 120
DB 61 CTTACGCGAAGCGCTAGCCATGGCGGTAGTAGTGTGCTAGACGCTCCAGGCC 120

QY 121 CCCCCTCCGGAGAGCCATAGTGTCTCGGAAACCGGTGAGTACACCGGAATTCGCGG 180
DB 121 CCCCCTCCGGAGAGCCATAGTGTCTCGGAAACCGGTGAGTACACCGGAATTCGCGG 180

```

```

181 AAGACTGGGTCTCTTTCTTGGATAAAACCACTCTATGCGCGGCAATTTGGGCGTGCCTCCG 240
DB 181 AAGACTGGGTCTCTTTCTTGGATAAAACCACTCTGTGCGCGGCAATTTGGGCGTGCCTCCG 240

241 CAAGACTGCTAGCCAGTAGCTTGGTTCGAAAGGCTTGTGTGTAATGCTGTGATGAGG 300
DB 241 CAAGACTGCTAGCCAGTAGCTTGGTTCGAAAGGCTTGTGTGTAATGCTGTGATGAGG 300

301 TGTCTTCGAGTGTCCCGGGGAGTCTCTGTAGACCGTGCACCATGAGCACAATCTCTAAAC 360
DB 301 TGTCTTCGAGTGTCCCGGGGAGTCTCTGTAGACCGTGCACCATGAGCACAATCTCTAAAC 360

361 TCAAAAGAAAAACAAAAGAAAAACAAACCGTGCACCAAGACGTTAAAGTTTCGCGGCGG 420
DB 361 TCAAAAGAAAAACAAAAGAAAAACAAACCGTGCACCAAGACGTTAAAGTTTCGCGGCGG 420

421 CGGCGAGATCGTGTGGCGGAGTATATCTTGTTCGCGGCGAGGGGCCCAAGTTCGGGTGTCG 480
DB 421 CGGCGAGATCGTGTGGCGGAGTATATCTTGTTCGCGGCGAGGGGCCCAAGTTCGGGTGTCG 480

481 CGCGACAAGGAGACTTCGGAGCGGTCCAGCCACGTTAGGAGCGCGCAGCCCATCCCTAA 540
DB 481 CGCGACAAGGAGACTTCGGAGCGGTCCAGCCACGTTAGGAGCGCGCAGCCCATCCCTAA 540

541 AGATCGGCGCTCCACTGGCAAAATCCTGGGAAAAACAGGATACCCCTGGCCCTCTATACG 600
DB 541 AGATCGGCGCTCCACTGGCAAGTCTCTGGGAAAAACAGGATACCCCTGGCCCTCTATACG 600

601 GAATGAGGAGTCTGGCTCGGCGAGGATGCTCTCTGTCGCGGAGGTTCCCGTCCCTCTTTG 660
DB 601 GAATGAGGAGTCTGGCTCGGCGAGGATGCTCTCTGTCGCGGAGGTTCCCGTCCCTCTTTG 660

661 GGGCCCCAATGACCCCCCGGCTAGTCTGGCGCAACGTTGGGTAAAGTCATGACACCCCTAAC 720
DB 661 GGGCCCCAATGACCCCCCGGCTAGTCTGGCGCAACGTTGGGTAAAGTCATGACACCCCTAAC 720

721 GTGCGGCTTTGCGGACCTCATGGGGTACATCCCTGCTGTGGGCGGCCCGCTCGGCGGCT 780
DB 721 GTGCGGCTTTGCGGACCTCATGGGGTACATCCCTGCTGTGGGCGGCCCGCTCGGCGGCT 780

781 GCGCAGAGTCTCTCGCGCATGCGGTGAGAGTCTCTGAGGACGCGGGTTAAATTTTGCACACG 840
DB 781 GCGCAGAGTCTCTCGCGCATGCGGTGAGAGTCTCTGAGGACGCGGGTTAAATTTTGCACACG 840

841 GAACTTACCGGTGCTCTCTTTCTATCTTCTGTGCGGCGCTGTCTCTGCTGCTGCTGCTGCT 900
DB 841 GAACTTACCGGTGCTCTCTTTCTATCTTCTGTGCGGCGCTGTCTCTGCTGCTGCTGCTGCT 900

901 CCGGCTCTCCGCTGCGGAAAGTGAAGAAACATCAGTACCGGCTACATGCTGCTGCTGCTGCTGCT 960
DB 901 CCGGCTCTCCGCTGCGGAAAGTGAAGAAACATCAGTACCGGCTACATGCTGCTGCTGCTGCTGCT 960

961 CACCAATGACAGCATTAATCCTGCGAGTCCAGGCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCT 1020
DB 961 TGCCTTACGAGCATCACTGCGAGTTCAGGCGGTTCAGGCGGTTCCTCCAGCTCCCTGCTGCTGCT 1020

1021 CCGGCTGCGGAAAGTGGGGAATGATCTCAGTGTGGAATACCGGTCTCACCGAATGTGCTG 1080
DB 1021 CCGGCTGCGGAAAGTGGGGAATGATCTCAGTGTGGAATACCGGTCTCACCGAATGTGCTG 1080

1081 CGTGAGCGGCGCGCGGCGCTCAGCAGGCGCTTGGCGACGACATCGACATGCTGCTGCTGCTGCT 1140
DB 1081 CGTGAGCGGCGCGCGGCGCTCAGCAGGCGCTTGGCGACGACATCGACATGCTGCTGCTGCTGCT 1140

1141 GTGCGGCGAGCGCGGCGCTCAGCAGGCGCTTGGCGACGACATCGACATGCTGCTGCTGCTGCTGCT 1200
DB 1141 GTGCGGCGAGCGCGGCGCTCAGCAGGCGCTTGGCGACGACATCGACATGCTGCTGCTGCTGCTGCT 1200

1201 AGCCCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
DB 1201 AGCCCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260

1261 CATCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320

```

```

RESULT 15
AF169002
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
AF169002
Hepatitis C virus isolate NDM228
polyprotein gene, complete cds.
AF169002.1
GI:6707279
Hepatitis C virus
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
1 (bases 1 to 9661)
Kurihara, C., Ishiyama, N., Nishiyama, Y., Fukushi, S., Kageyama, T.,
Katayama, K. and Miura, S.
Molecular characterization of hepatitis C virus genotype 2a from
the entire sequences of four isolates
J. Med. Virol. 64 (4), 466-475 (2001)
21361470
11468731
2 (bases 1 to 9661)
Katayama, K. and Kurihara, C.
Direct Submission
Submitted (14-JUL-1999) Basic Research Division, BioMedical
Laboratories, Inc., 1361-1 Matoba, Kawagoe, Saitama 350-1101, Japan
Location/Qualifiers
1..9661
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/isolate="NDM228"
/db_xref="taxon:11103"
/note="genotype: 2a"
341..9442
/codon_start=1
/product="polyprotein"
/protein_id="AAF25610.1"
CDS

```


3661	TGCCAAGAGGCACTTGGTAGGGTGGCCAGTCTCCCGGACCAAAATCTTTGAGCCGTG	3720	QY
3721	CAGTGTGAGCCGTCGACCTGTACTGTGTCAAGCGGACGCTGATGTATCCCGGCTCG	3780	Db
3721	CAGTGTGAGCCGTCGACCTGTACTGTGTCAAGCGGACGCTGATGTATCCCGGCTCG	3780	Db
3781	AAGACGGGGGCAAAACGGGGAGCGGTACTCTCTCCCGAGACCTCTTTCCACCTTGAAGGG	3840	QY
3781	AAGACGGGGGCAAAACGGGGAGCGGTACTCTCTCCCGAGGCTCTTTCCGACCTGAAGGG	3840	Db
3841	GTCTCAGGAGGCCCGGTGTATATGCCCGAGGGCCACGCTGTGCGAGTCTTCCGGGAGC	3900	QY
3841	ATCATCGGGGGACCGGTGTCTGTCTGTAGGGGCCACGCTGTGCGGATCTTCCGGGAGC	3900	Db
3901	TGTGTCTCTCGGGGCTGCTAAGTCCATAGATTTTATCCCGGTTGAGACATCTGCACAT	3960	QY
3901	TGTGTCTCTCGGGGCTGCTAAGTCCATAGATTTTATCCCGGTTGAGGCGCTGCACAT	3960	Db
3961	CGTCACGCGGTCCCGACCTTTAGTGAACACAGCACACCACTGTGTGCCCGAGACCTTA	4020	QY
3961	CGTCACGCGGTCTCCACCTTTAGCGACACACAGCACACCACTGTGTGCCCGAGACCTTA	4020	Db
4021	TCAGGTGGGTACTGTGATGCCCGGACTGCGAGTGGAAAGAGCACCAAGTTCTGTGCG	4080	QY
4021	TCAGGTGGGTACTGTGATGCCCGGACTGCGAGTGGAAAGAGCACCAAGTTCTGTGCG	4080	Db
4081	ATATGTCTCAGGGGTATAAGTGTAGTGTCTTAATCCCTCAGTGGCTGCCACCTGGG	4140	QY
4081	ATATGTCTCAGGGGTATAAGTGTAGTGTCTTAATCCCTCAGTGGCTGCCACCTGGG	4140	Db
4141	GTTTGGGGCTACTGTCTAAGGCACATGGCATCAATCCCAACATTAGGACTGGAGTCAG	4200	QY
4141	GTTTGGGGCTACTGTCTAAGGCACATGGCATCAATCCCAACATTAGGACTGGAGTCAG	4200	Db
4201	GACTGTGACGACCGGGGCGCCATFACGTAATCTTCAATATGACCAATTCCTCGCGATGG	4260	QY
4201	AACTGTGACGACCGGGGCGCCATFACGTAATCTTCAATATGACCAATTCCTCGCGATGG	4260	Db
4261	GGGCTGTGCGGGGCGGCTACGACATCATATGATGATGATGATGATGATGATGATGATG	4320	QY
4261	GGGCTGTGCGGGGCGGCTACGACATCATATGATGATGATGATGATGATGATGATGATG	4320	Db
4321	TACCACCATCTTGGCATCGGACAGTCTTGTGATCAAGCAGACAGACAGTGGGGTCAGACT	4380	QY
4321	TACCACCATCTTGGCATCGGACAGTCTTGTGATCAAGCAGACAGACAGTGGGGTCAGACT	4380	Db
4381	AACTGTCTGGCTACAGTACGCGCCCTGGGTGAGTGAACACCCCGGACCAACATAGA	4440	QY
4381	CACTGTACTGGCCACGCGCCACGCGCCCGGGTGGTGAACACTCCCATCCCAACATAGA	4440	Db
4441	GGAGTGGCCCTTGGGAGGAGGCGGAGATCCCTCTATGGGAGGGGATTCCTCCGTGC	4500	QY
4441	GGAGTGGCCCTTGGGAGGAGGCGGAGATCCCTCTATGGGAGGGGATTCCTCCGTGC	4500	Db
4501	TTACATCAAGGGAGGAGACATCTGATCTTCTGCTTCAAGAAAAAGTGTGACGAGCT	4560	QY
4501	TTACATCAAGGGAGGAGACATCTGATCTTCTGCTTCAAGAAAAAGTGTGACGAGCT	4560	Db
4561	CGCGGCGCCCTTGGGGGTATGGGCTTGAATCTGATGGGAGGAGTGGAGCT	4620	QY
4561	CGCGGCGCCCTTGGGGGTATGGGCTTGAATCTGATGGGAGGAGTGGAGCT	4620	Db
4621	CTCCGTAATACCAACTCAGGAGAGTGTGTGCTGCGCACCGACCCCTCATGACAGG	4680	QY
4621	CTCCGTAATACCAACTCAGGAGAGTGTGTGCTGCGCACCGACCCCTCATGACAGG	4680	Db
4681	GTATACTGGGACTTTGACTCCGTGATCGATGCAAGAGTGGGCTCACTCAAGTTGTAGA	4740	QY
4681	GTTCGCTGGAGACTTTGACTCTGTGTGATGCACTGCAAGTGGGCTCACTCAAGTGTAGA	4740	Db
4741	CTTCAGTTTAGACCCCACTTACCAATAACACACAGATGTGCTCCCTCAAGACGCTGTCTC	4800	QY
4741	CTTCAGTTTAGACCCCACTTACCAATAACACACAGATGTGCTCCCTCAAGACGCTGTCTC	4800	Db
4801	ACGTAGCCAGCCCGGGTTCGCAAGGTAGGGGAAGACTGGGCAATTTATAGGTATGTTTC	4860	QY
4801	ACGTAGCCAGCCCGGGTTCGCAAGGTAGGGGAAGACTGGGCAATTTATAGGTATGTTTC	4860	Db
4861	CACTGTGTAGCGAGCTCTCAGGAATGTTTGAAGTGTAGTGTCTCTGTAGTGTCTACAGCG	4920	QY
4861	CACTGTGTAGCGAGCTCTCAGGAATGTTTGAAGTGTAGTGTCTCTGTAGTGTCTACAGCG	4920	Db
4921	AGGGCCCGCATCGTATGAGTCAACCATCGAGACCAAGCTCAGGCTCAGGGCGTATTT	4980	QY
4921	AGGGCCCGCATCGTATGAGTCAACCATCGAGACCAAGCTCAGGCTCAGGGCGTATTT	4980	Db
4981	CAACACCGCCGGTGTGCTGTGTGCCAAGACCATCTTGAAGTGTGGAGGCGAGTTTTCAC	5040	QY
4981	CAATACCGCTGTGCTGTGTGCCAAGACCATCTTGAAGTGTGGAGGCGAGTTTTCAC	5040	Db
5041	CGGCTCTACACATAGATGCCCATCTTCTTTCCAGACAAACAAATCGGGGGAGAAATTT	5100	QY
5041	CGGCTCTACACATAGATGCCCATCTTCTTTCCAGACAAACAAATCGGGGGAGAAATTT	5100	Db
5101	CGCATCTTAAACAGCCTACAGGCTACAGGTACAGTGTGCGCTAGGGCCAAAGCCCCCGGTC	5160	QY
5101	CGGTACTTAGTAGCTTATCAGGCCACAGTGTGCGCTAGAGCCAAAGCCCCCGGTC	5160	Db
5161	CTGGGACGTCATGTGGAAGTGTGTGACTCGACTCAAGCCCCACACTCGTGGGCCCAACACC	5220	QY
5161	TTGGGACGTCATGTGGAAGTGTGTGACTCGACTCAAGCCCCACACTCGTGGGCCCAACACC	5220	Db
5221	TCTCTGTATCCGCTTGGGCTCTGTGTACAAAGAGGTCACTTCCACATCTCCGTGAGCA	5280	QY
5221	TCTCTGTATCCGCTTGGGCTCTGTGTACAAAGAGGTCACTTCCACATCTCCGTGAGCA	5280	Db
5281	ATACATCGCCACCTGATGCAAGCCGACCTTGAAGTCAATGAGGTCATGAGGCTCTGGC	5340	QY
5281	ATACATCGCCACCTGATGCAAGCCGACCTTGAAGTCAATGAGGTCATGAGGCTCTGGC	5340	Db
5341	AGGGGAGTCTTGGCGCGCTCGCCGCTATTTGCTTGGCGACCGGGTGTGTTGCAATCAT	5400	QY
5341	TGGGGAGTCTTGGCGCGCTCGCCGCTATTTGCTTGGCGACCGGGTGTGTTGCAATCAT	5400	Db
5401	CGGCGCTTGCACATTAACAGGAGCGCTGTGTGGCGCGGACAGGAGTCTCTATAGA	5460	QY
5401	CGGCGCTTGCACATTAACAGGAGCGCTGTGTGGCGCGGACAGGAGTCTCTATAGA	5460	Db
5461	GGCTTTGTATGATGAGTGGAGAAATGTGCTTGAAGTCAAGGCTTATTTGCAAGAGGCGAG	5520	QY
5461	GGCTTTGTATGATGAGTGGAGAAATGTGCTTGAAGTCAAGGCTTATTTGCAAGAGGCGAG	5520	Db
5521	GATAGCCGAGATGCTGAAAGTCCAAGATCCAAGGCTTATTTGCAAGAGTCTTCCAAACAGC	5580	QY
5521	GATAGCCGAGATGCTGAAAGTCCAAGATCCAAGGCTTATTTGCAAGAGTCTTCCAAACAGC	5580	Db
5581	TCAAGCATACAAACCCATCTGTGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTT	5640	QY
5581	CCAGGACATACAAACCCATCTGTGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTT	5640	Db
5641	ACACATGTGAACTTCAATAGCGGCTTCAATACCTCGAGGAGTCTCAACATCTGCCAGG	5700	QY
5641	ACACATGTGAACTTCAATAGCGGCTTCAATACCTCGAGGAGTCTCAACATCTGCCAGG	5700	Db
5701	GAACCTTGCAGTGTGCTTCCATGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT	5760	QY
5701	GAACCTTGCAGTGTGCTTCCATGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT	5760	Db
5761	AAGACCATCTATCTTCTCAACATTTTGGGGGCTGCTAGCATCTCCAAATTTGCACACCC	5820	QY
5761	TAGCACCATCTATCTTCTCAACATTTTGGGGGCTGCTAGCATCTCCAAATTTGCACACCC	5820	Db
5821	CGGGGGGCGGCTTGGCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT	5880	QY
5821	CGGGGGGCGGCTTGGCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT	5880	Db

Copied from 09540843 on 05/19/2004

PFLLILLALPOQAYADYASVHGOIGAAILLILITLFTLTPGYKTLRSLSLWMLCYLLTL
 GYRAVOWEAPPMQARGDGIITWAATILCPGVFDTITKLLAVLTPWLLRDLALTSVP
 YFVHALLCMTWVHAGGAYVOMALLALGWSTYIYDHLTPMSDWAASGLRDLA
 VAVEPIIFSPMEKVIWGAETAAAGDILYGLPVSARLGREILLGPTGYTSKGRLL
 APITAYQWTRGLLGAIVVSMGRDKTBQAGEIQLVSTQSFGLTSISGVWTVYHG
 TNGKTLASRGFPVTQMSAAGDLVWSPGPGTKSLPCTCGAVDLVLYTRNADVIPA
 RRRGDKGALLSPRLSTLTKSGSGVPLCPRGHAGVIFRAAVCSRGVAKSIDIFIPVEA
 LDIIVTRSPFTSONSTPAVQTYQVGLHAPTSGSKSTKVPVAYAAQYKVLVLPNSV
 AATLPGFASLKAHGINPNIIRTCVTRVTTGGPIITYSGKFLADGDCSGCAVDIICD
 ECHAVDATTILGIVTLDQAEATAGVLTATVATATPGSVTTPHNIETALGOEIEIP
 FYGRALPILSHIKGGRHLIFCHSKKCDLAAALRMGLNNAVAYYRGLDVSIIPTQGDV
 VVATDALMTGAGDFDSDVIDCNVAVTQVDFELDPFTITQTVQDDAVSRQRGR
 TGRGLRTYRYSTGSRASGDFSVLCECYDAGAAWYBLTPTVETVIRLGAFTNPGL
 PVQDHLFEWAVFTGLTHIDAFHSQTKSGENFAYLVAQYATVCARAKAPPNDV
 MWKCLTRLAPTGLVTPPLLYRLGVSNTVEVLTHTPKYLATQADLEVTMTWTLVAG
 GYLAAVAAYLACVSIIGRLHINORAVAPDKVELYEADEMEECASRAALIEEGQ
 RMAEMUKSKIGLLOQASQAQDIQPAVOASPKVQFQWAKHWNFIQIYQVLAGLST
 LPGNPAVAMFSAALTSPLSTSTLILNIGLWLASQIAPPAGATGFGVSLGVAA
 VSGVGLKVLVDILAGYGTIGSALVAFKIMSEKPEMEDVNVNLLPGLISGLALVVG
 ICATILRRHVGPGEAVQWNRILIAFASINHVAPHYVTVESDASORVQLGLSLTIT
 SLILRLHNTIDPCIPPCAGSLMDMDVCTILTDPKNWLTSKLPFKMPGLPFTSCQ
 KYGVWAGTGIMTRPCGASISGNVRLGSMRLTGPCTCMNIWOCTFPCINCYTSQC
 MPKAPNFKTAIRWAASTAEVITQHPYSIITGLTIDNLKVPQCPSPBFFSWDGV
 QHIFAPTRFPPFRDEVSFCVGLNSFVGSQPCDPEDPDVLMGLMDTDFSHITABAA
 ARRLARGSPSEASSASQISAPLSRATCTTHGTYDMDVMEANLPMGDVTRIESES
 KVVVLDLSMVBERDLRPSIPSEYMLPKTPPALPAWAPDYNPPLVPSGWKRPDY
 OPTVAGCALPPKTPPTPPKRTVGLSESTIGDALOOLAKTFPGOPPSPGDSGLS
 ARSKYGFAGKVESLGRVNIHKSVMKDLSDSOTPIPTIIMAKNEVFCVDPKTGGK
 KAARLLIVPDGLVRCVKEALYDOKLPOAVMGASYQFYSPAQRVEFLKLAWADKR
 EPMGFSDVTRCPDSTERTDIERESIYOACSLPEBAHTAIHSLTERLVYGGPMLNSK
 GQACGYRRCBASGLVITSMGNTITCYKALAAKAGIIVAPTMVLCGDLDIVSESQG
 TEFEDERLRAFTAMTRYAPPDPRDEYDLELITSCSNVSVAGLPQOGRRYLYTR
 DPTTIPARAWEVTRHPSNVNIGNIIQVAPTIWVRMLTHFFSILMAQDITLDQNLN
 FEMYGVSYPVLDLPAIIBERHLGDAFSLHTYSPHELTRVASALRKLGLAPLRAWS
 RARAVRASLISRGRAAICORVLPFNWAVTKLKLTLPLPEARLIDLSWFTVGAGGDI
 YHVSRRARPRELLILLLLVGVGLALLPAR"

BASE COUNT 1985 a 2828 c 2693 g 2185 t

ORIGIN
 Query Match 86.0%; Score 8349; DB 14; Length 9691;
 Best Local Similarity 91.4%; Pred. No. 0;
 Matches 8871; Conservative 0; Mismatches 820; Indels 20; Gaps 1;
 1 ACCCGCCCCCTAATAGGGGGGACACTCGGCATGAATCACTCCCTGTGAGGAACACTGT 60
 1 ACCCGCCCCCTAATAGGGGGGACACTCGGCATGAATCACTCCCTGTGAGGAACACTGT 60
 61 CTTTCAGCAGAAAGCGCTAGCGATGCGGTAGTATGAGTGTGTCAGAGCTCCAGGCC 120
 61 CTTTCAGCAGAAAGCGCTAGCGATGCGGTAGTATGAGTGTGTCAGAGCTCCAGGCC 120
 121 CCCCCTCCCGGAGAGCCATAGTGGTCTCGGGAAACCGGTGAGTACACCGGAATTCGCGG 180
 121 CCCCCTCCCGGAGAGCCATAGTGGTCTCGGGAAACCGGTGAGTACACCGGAATTCGCGG 180
 181 AAGACTGGGTCTTTCTTGGATAAACCCACTCTATGCCCCGCCATTTGGGCGTCCGCCCG 240
 181 AAGACTGGGTCTTTCTTGGATAAACCCACTCTATGCCCCGCCATTTGGGCGTCCGCCCG 240
 241 CAAGACTGCTAGCGGAGTGGTGGGTGGGAAAGCGCTTGTGGTACTGCTGATAGG 300
 241 CAAGACTGCTAGCGGAGTGGTGGGTGGGAAAGCGCTTGTGGTACTGCTGATAGG 300
 301 TGCTTGGAGTGGCCCCGGAGGTCTCGTAGACCGGTGACCATGAGCACAATCTCTAAACC 360
 301 TGCTTGGAGTGGCCCCGGAGGTCTCGTAGACCGGTGACCATGAGCACAATCTCTAAACC 360
 361 TCAAGAAAAACCAAAAGAAACACCAACCGTCGCCCAACAGAGCTTTAAGTTTCGGGCGG 420
 361 TCAAGAAAAACCAAAAGAAACACCAACCGTCGCCCAACAGAGCTTTAAGTTTCGGGCGG 420

Qy 421 CGGCCAGATCGTTGGCGGAGTATCTTGTTCGCGCAGGGGCCAGGTTGGGTGTCG 480
 Db 421 CGGCCAGATCGTTGGCGGAGTATCTTGTTCGCGCAGGGGCCAGGTTGGGTGTCG 480
 Qy 481 CGCGACAGGAAGACTTCGGAGCGGTCCAGGCCAGTGGAGGCGCAGCCATCCCTAA 540
 Db 481 CGCGACAGGAAGACTTCGGAGCGGTCCAGGCCAGTGGAGGCGCAGCCATCCCTAA 540
 Qy 541 AGATCGGCGCTCCACTGGCAATCTGGGGAACACAGGATACCCCTGCGCCCTATACGG 600
 Db 541 AGATCGGCGCTCCACTGGCAAGTCTTGGGGAACACAGGATACCCCTGCGCCCTATACGG 600
 Qy 601 GAATGAGGACTCGGCTGGGAGGATGCTCTGTCCTCCCGGAGGTTCCCGTCCCTCTTG 660
 Db 601 GAATGAGGCGCTCGGCTGGGAGGATGCTCTGTCCTCTGTCCTCCCGGAGGTTCCCGTCCCTCTTG 660
 Qy 661 GGGCCCCAATGACCCCGGATAGTCCGCGAAGTGGTGAAGTTCATCGATACCCCTAAC 720
 Db 661 GGGCCCCAATGACCCCGGATAGTCCGCGAAGTGGTGAAGTTCATCGATACCCCTAAC 720
 Qy 721 GTGCGGCTTTGCGGACCTCATCGGCTTACATCCCTGTCGTGGGCGCCCGCTCGGCGCGT 780
 Db 721 GTGCGGCTTTGCGGACCTCATCGGCTTACATCCCTGTCGTGGGCGCCCGCTCGGCGCGT 780
 Qy 781 CGCCAGAGCTCTCGGCGCATGGGCTGAGAGTCTCTGGAGGACGGGTTAAATTTTGCAACAG 840
 Db 781 TGCAGAGCTCTCGGCGCATGGGCTGAGAGTCTCTGGAGGACGGGTTAAATTTTGCAACAG 840
 Qy 841 GAACTTACCGGTTGCTCTTTCTATCTTCTTGTGGGCGCTGCTGCTGATCACCAC 900
 Db 841 GAACTTACCGGTTGCTCTTTCTATCTTCTTGTGGGCGCTGCTGCTGATCACCAC 900
 Qy 901 CCGGCTCTCGGCTGCGGAGTGAAGAACATCAGTACCGGCTACATGCTGCTAAACGACTG 960
 Db 901 TCGGCTCTCTGCTGCTGAAGTGAAGAACATCAGGACGACTACATGCTGCTAAACGACTG 960
 Qy 961 CACCAATGACAGCATTTACTGCGAGCTCCAGGCTGCTGCTCCAGCTCCCGGGTGGCT 1020
 Db 961 TTTCAATAACAGCATCACCTGCGAGCTTGGGCTGAGTCTTCCAGCTCCCGGGTGGCT 1020
 Qy 1021 CCGGTCGGAAGAGTGGGGAATGTCATCTCAGTCTGATACCGTCTCAGCAATGTCG 1080
 Db 1021 CCGGTCGGAAGAGTGGGGAATGTCATCTCAGTCTGATACCGTCTCAGCAATGTCG 1080
 Qy 1081 CGTGACGCGCGCGGCGCTTCAACGAGGCTTTCGAGGACGACATGATGCTGTTGAT 1140
 Db 1081 TGTCCGCGAGCGCGGCGCTTCAACGAGGCTTTCGAGGACGACATGATGCTGTTGAT 1140
 Qy 1141 GTCCGCGACGCTGCTGCTGCTGCTTCACTGAGGAGCTTTCGAGGAGCTGATGCTGCG 1200
 Db 1141 GTCCGCGACGCTTGTCTGCTGCTTCCAGCTGAGGAGCTTTCGAGGAGCTGATGCTGCG 1200
 Qy 1201 AGCCCAATTTGTCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
 Db 1201 AGCCCAATTTGTCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
 Qy 1261 CATCTACCTTGTGATCCTTGTGATCCTTGTGATCCTTGTGATCCTTGTGATCCTTGTGAT 1320
 Db 1261 CATTTACCTTGTGATCCTTGTGATCCTTGTGATCCTTGTGATCCTTGTGATCCTTGTGAT 1320
 Qy 1321 GCGCAGGCTTACCATGATCTTGGCGTACGCGATGCGGTGCTGCTGCTGCTGCTGCTGCTGCT 1380
 Db 1321 GCGCAGGCTTACCATGATCTTGGCGTACGCGATGCGGTGCTGCTGCTGCTGCTGCTGCTGCT 1380
 Qy 1381 CATTAGCGGCTTACCATGATCTTGGCGTACGCGATGCGGTGCTGCTGCTGCTGCTGCTGCTGCT 1440
 Db 1381 TATTAGCGGCTTACCATGATCTTGGCGTACGCGATGCGGTGCTGCTGCTGCTGCTGCTGCTGCT 1440
 Qy 1441 GTGCGGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
 Db 1441 GTGCGGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500

Copied from 09540843 on 05/19/2004

GSTLGLKVLVDILAGYAGISGALVAFKIMSGEKSPMEDVNLPLGILSPGALVWGI
 CAAIRRHVTPEDGCAVWMNRILJAFASGNHVAFTVTEDESASORVOLLGLSLITSC
 LLRLNHWITTEPCIPICAGSLRDLVDWVCSILITDFKNMLTSLKPKMPLGLSITSCOK
 GYKVGWAGTIGMTTRCPGNAVISGNVRLSGMRITGPKTQNTWQGTFFPINCYTEGCV
 PKPAPFKTAIRVAASEYAEVATQHGYSITITGLITDNLKVPCLPSPFPFFSWDGVQ
 IHRFAPTPKPPFFRDEYSFCVGLNSFVVSQPCDPEPDTDLVMSLITDPSHITAEASK
 RLLAGSPFSEASLSAASLSRATCTHGTKYDVMVDWADLFLNKGVDVTRIESESK
 VVLDYLDPMVBESTDREPSIPSEYMLPRNRPALPAWAPDYNPLPVSWKRPDYQ
 PPTVACGLPPKPTTPPRRERTVGLSESTIGDALOOLAVKSGFQPPSPGSDGSLST
 GADAASGGTSPDELALSETGISISIPPLEBEPGPDLEPEVELOPPPPQGEVAPG
 SDGWSSTCEEDAVVCCMSYSYSGALITPCSPBEELPINPLNSLLRLHKNVYC
 TTSKASLAKRKTFTFORMQVLDYDVLKDIKLAASKVSAARLLTLEECQLTTPHSA
 RSKYGFKAERLSTRANVHIKSWMEDLLESQTPITIMAKNEVFCVPTKGGKK
 AARLIYVPLDGVRCMKMALYDVKLOPAVMGASGYFOVSQRPVEFLLLKAWAEKKG
 PMGFSYDTRFSTFOTVERDIKTEESIYQASLPEEARTAIHSILTERLYVGGPMPNSG
 QSCGYRRCASGVLTSMGNTICYVKALAAKAGIIVAPTMVLVCGDDLVIVISEOGT
 EEDERNIAFTEAMTYSAPPDGNPRPEYDELITSCSNVSVALGPPQRRRYLYTRD
 PATPIARAARETRSPVNSJUNIIQYAPTIVWRMLTHFFPSILMAQDTLDONLNF
 EMYGSVYSPLDLPAIIBRLHGLGAFSLHTVPHELTRVAGSALRKLGAAPLRTWKS
 ARAVRASLISRGRAAVCGRYLPFNWVRKLLITLPLEARLLDLSSWFTVGAGGGDIY
 HVSRRARPRLLLSLILFVGLFLLPAR"
 BASE COUNT 1992 a 2850 c 2678 g 2163 t
 ORIGIN

Query Match 86.1%; Score 8364.2; DB 14; Length 9683;
 Best Local Similarity 91.6%; Pred. No. 0;
 Matches 8895; Conservative 0; Mismatches 788; Indels 28; Gaps 3;
 1 ACCGGCCCTAATAGGGGCGACATCGCGCATGAATCACTCCCTGTGAGGAACACTACTGT 60
 1 ACCGGCCCTAATAGGGGCGACATCGCGCATGAATCACTCCCTGTGAGGAACACTACTGT 60
 61 CTTTCAGCGAAGAGCGTAGCGCATGCGTTCAGTATGAGTGTGTCAGCGCTCCAGGCC 120
 61 CTTTCAGCGAAGAGCGTAGCGCATGCGTTCAGTATGAGTGTGTCAGCGCTCCAGGCC 120
 121 CCCCCTCCCGGAGAGCCATAGTGGTCTCGGGAACCGGTGAGTACACCGGAATTCGCGG 180
 121 CCCCCTCCCGGAGAGCCATAGTGGTCTCGGGAACCGGTGAGTACACCGGAATTCGCGG 180
 181 AAGACTGGGTCTCTTTCTTGGATAAACCCACTATATGCCCGGTCAATTTGGCGGTGCCCCG 240
 181 AAGACTGGGTCTCTTTCTTGGATAAACCCACTATATGCCCGGTCAATTTGGCGGTGCCCCG 240
 241 CAGACTGTCTAGCGAGTACGTTGGTTCGGAAGCGCTTGTGGTACTGCTGTATGAGG 300
 241 CAGACTGTCTAGCGAGTACGTTGGTTCGGAAGCGCTTGTGGTACTGCTGTATGAGG 300
 301 TGCTTCGAGTGTCCCGGAGGTCTCGTAGACCGTGCACCATGAGCACAAATCCTAAAC 360
 301 TGCTTCGAGTGTCCCGGAGGTCTCGTAGACCGTGCACCATGAGCACAAATCCTAAAC 360
 361 TCAAGAAACCAAAAGAAACCAACCGTTCGCCCAAGAGCTTAAAGTTTCGCGGCGG 420
 361 TCAAGAAACCAAAAGAAACCAACCGTTCGCCCAAGAGCTTAAAGTTTCGCGGCGG 420
 421 CGGCCAGATGTTGGCGGAGTATCTTGTTCGCGCAGGCGCCAGGTGGGTGCG 480
 421 CGGCCAGATGTTGGCGGAGTATCTTGTTCGCGCAGGCGCCAGGTGGGTGCG 480
 481 CGGCACAGGAAGACTTCGAGCGGTCCAGCACCGTGGAAAGCGCCAGCCATCCCTAA 540
 481 CGGCACAGGAAGACTTCGAGCGGTCCAGCACCGTGGAAAGCGCCAGCCATCCCTAA 540
 541 AGATCGGCGTCCATCGGCAATTCCTGGGAAACCAAGATACCCCTGCGCCCTATACGG 600
 541 AGATCGGCGTCCATCGGCAATTCCTGGGAAACCAAGATACCCCTGCGCCCTATATGG 600
 601 GAATGAGGACTCGGCTGGGAGGATGCTTCCCTGCGCCCGAGGTTCGGTCCCTTTG 660
 601 GAACGAGGACTCGGCTGGGAGGATGCTTCCCTGCGCCCGAGGTTCGGTCCCTTTG 660
 661 GGGCCCCAATGACCCCGGATAGGTTCGCGCAACGTGGGTGAGGTATCGATACCCCTAAC 720

661 GGGCCCCAATGACCCCGGATAGGTTCGCGCAACGTGGGTGAGGTATCGATACCCCTAAC 720
 721 GTGCGGCTTTGGCGGACTCATGGGTATCATCTCTGTCTGGGGCCCGCTCGGCGCGT 780
 721 GTGCGGCTTTGGCGGACTCATGGGTATCATCTCTGTCTGGGGCCCGCTCGGCGCGT 780
 781 CGCCAGAGCTCTCGCGCATGGGTGAGAGTCTCTGGAGGACGGGTAAATTTGCAACAGG 840
 781 TGCCAGAGCTCTCGCGCAGCGGTGAGAGTCTCTGGAGGACGGGTAAATTTGCAACAGG 840
 841 GAACTTTACCCGGTGTCTCTTTTCTATCTTTTGTGTGGCCCTGCTCTCTGTCATCAC 900
 841 GAACTTTACCCGGTGTCTCTTTTCTATCTTTTGTGTGGCCCTGCTCTCTGTCATCAC 900
 901 CCGGCTCTCGCTGCGGAGTGAAGAAACATCATAGTACCGGCTCATAGGTGACTAAACAG 960
 901 CCGGCTCTCTGCTGCGGAGTGAAGAAACATCATAGTACCGGCTCATAGGTGACTAAACAG 960
 961 CACCAATGACAGCATTTACCTGCGAGCTCCAGGCTGCTCTCTCCAGCTCCCGGGTGG 1020
 961 TTTCTAATGACAGCATCACCTGCGAGCTTCAGGCGCGGCGCTCCAGCTCCCGGGTGG 1020
 1021 CCGGTCGCGAAGTGGGAATGCAATCTCAGTGTGATACCGGCTTCACCGAATGTGGC 1080
 1021 CCGGTCGAGAGGGTGGGAACACATCATCGGTGTGGATACCGGCTTCACCGAATGTGGC 1080
 1081 CGTGACGCGCGCGGCGGCTCATCGAGGCTTGGCGGACGACATGACATGGTGTGAT 1140
 1081 TGTGCGGAGCGCTGCGGCGCTCACGAGGCTTGGCGGACGACATGACATGGTGTGAT 1140
 1141 GTCCGCGAGCTGCTGCTGCGCTCTCAGTGGGAGCTCTCGGTTGGGCTGATGCTCGC 1200
 1141 GTCCGCGAGCTGCTGCTGCGCTCTCAGTGGGAGCTCTCGGTTGGGCTGATGCTCGC 1200
 1201 AGCCCAATTTGTTCTGCTGCGCGACCACTCTGTTGTTGTCAGAGCTGCAATTTGCTC 1260
 1201 ASCTCAGATGTTTCAATTAATCTGCGCGACGACCACTGTTGTTGTCAGGAATGCAATTTGCTC 1260
 1261 CATCTACCTGTGTATCATCATCTGGAACCGCATGAGGACATGATGATGATGATGATG 1320
 1261 CATCTACCTGTGTATCATCATCTGGAACCGCATGAGGACATGATGATGATGATGATG 1320
 1321 GCCACGCTACCATGATCTTGGCGTACGCGATGCGTGTGCGGAGCTCATTTATGACAT 1380
 1321 GCCACACCAATGATCTTGGCGTACGCGATGCGTGTGCGGAGCTCATTTATGACAT 1380
 1381 CATTAGCGGGCTCATTTGGGCGCTCATGTTGCGGCTTGGCTTCTTCTATCAGGAGC 1440
 1381 CATCGCGGGCTCATTTGGGCGCTCATGTTGCGGCTTGGCTTCTTCTATCAGGAGC 1440
 1441 GTGGGCGAAGTGTGTTGTCATCTTGTGCGCGCGGGTGGAGCGCGGACCATAC 1500
 1441 GTGGGCGAAGTGTGTTGTCATCTTGTGCGCGCTGCGGCTGGAGCGGACACACAG 1500
 1501 TGTTCGGGCTTCTGCGCGGACACACCGGCGCTCACAGCTTATTTGACATGGGCCC 1560
 1501 --- TGGGAGTCTGTCGGGTACGCCACCTTCGGCATGTCGGCTTCTTCACTCTGCTG 1557
 1561 CAGGCGAAGAAATCCAGCTCTGTTAAACCAATGGAGCTGGGACATCAACCGGCGGCT 1620
 1558 TAAGCAGAACATTTAGCTCATTTAAACCAATGGAGTGGGACATCAACCGGCTGCT 1617
 1621 GAACTGCAATGACTCTTGTGACACCGGCTTATCGGCTCTCTGTTCTACACCGAGCT 1680
 1618 GAACTGCAATGACTCTTGTGACACCGGCTTATCGGCTCTCTGTTCTACCGGCT 1677
 1681 CAACTGCTCAGGATGTCGCGAAGCATGTCGCGCTGCGGCTGCGGCTTCCGGGT 1740
 1678 CAACTGCGACAGATGTCGCGGCTGTCGCGCTGTCGCGCTGCGGCTTCCGGGT 1737
 1741 GGGTGGGCGCTTGTGCAATATGAGGATTAATGTGCAATCAATCAAGGATATGAGACCT 1800

7501	TTCTACATCTCTCATGCCCCCTTCAGGGGGAGCCTTGGGGATCCAGACTCGAGGCTCTGA	7560
7561	GCAGGTAGAGCCCCCAACCCCCCCCCACGGGGGGGTGGCAGCTCCCGGCTCGACATCGGG	7620
7561	GCAGGTAGAGCTTCAGCCTCCCCCTCAGAGGGGGGASGTAGCTCCCGGCTCGACATCGGG	7620
7621	GTCTCGTCTTACTTGCTCCGAGGAGATGATCTCCGTGCTGTGCTCTCCATGTCTACTCTC	7680
7681	CTGACCGGGGCTCTAATAATCTCTGTAGCTCCGAAAGAGAGNAGTTACCGATTAAACC	7740
7681	CTGACCGGGGCTCTAATAAATCTCTGTAGCTCCGAAAGAGAGNAGTTACCGATTAAACC	7740
7741	CTTGAGCAATCTCCCTGTTGGCATATCACAAAGGTGTACTGTACCAACAACAAGAGCGC	7800
7741	CTTGAGCAATCTCGTGTTCGGATACCAACAAGGTGTACTGTACCAACAACAAGAGCGC	7800
7801	CTCACTAAGGGCTTAAAGGTTAACTTTTGATAGGATCAAAGTGTCTGCACTCTCATACGA	7860
7801	CTCACTAAGGGCTTAAAGGTTAACTTTTGATAGGATCAAAGTGTCTGCACTCTCATACGA	7860
7861	CTCAGTCTTAAAGGACATTAAGCTTAGCGGCTTCAAAGTTCACCGCAAGGCTCCTCACCAT	7920
7861	CTCAGTCTTAAAGGACATTAAGCTTAGCGGCTTCAAAGTTCACCGCAAGGCTCCTCACCAT	7920
7921	GGAGGAGGCTTGCGAGTTAAACCCACCCATCTCTGCAAGATCTATAATATGGTTCCTGGGC	7980
7921	GGAGGAGGCTTGCGAGTTAAACCCACCCATCTCTGCAAGATCTATAATATGGTTCCTGGGC	7980
7981	TAAAGAGGTCCGAGCTTGCTCCGGAGGGCGCTTAAACCATCAAGTCCGCTGTGGAAGA	8040
7981	TAAAGAGGTCCGAGCTTGCTCCGGAGGGCGCTTAAACCATCAAGTCCGCTGTGGAAGA	8040
8041	CTCTCTGGAGACTCAGAAACACCAATTTCCCAACCATTAATGGCCAAAATGAGGTGTT	8100
8041	CTCTCTGGAGACTCAGGAACACCAATTTCTTACGACCATAATGGCCAAAATGAGGTGTT	8100
8101	CTGCTGGACCCCAACAGGGGGGCAAGAAAGAGCTCGCTTATGTTTACCTCCACCT	8160
8101	CTGCTGGAAACCAACAGGGGGGTAGAAAGAGCTCGCTTATGTTTACCTCCACCT	8160
8161	CGCGTCAAGGCTCTCGAAGATGGCCCTTTATGACATTACACAAAACCTTCTCTCAGGC	8220
8161	CGCGTCAAGGCTCTCGAAGATGGCCCTTTATGATGTACACAAAAGCTTCTCTCAGGC	8220
8221	GGTGATGGGGCTTCTTATGGATTCCAGTATTTCCCGCTCAGGGGTAGAGTTCTCTT	8280
8221	GGTGATGGGGGCTTCTTATGGCTTCCAGTACTCCGCTCGCCCAACGGGTGAGTTCTCTT	8280
8281	GAAGGATGGGGGAAAGAGGACCTATGSGTTTTTCTGTATGATACCCGATCTTTTGA	8340
8281	GAAGGATGGGGGATAGAAAGACCTATGGTTTTTCTGTATGATACCCGATCTTTTGA	8340
8341	CTCAACCGTCACTGAGAGAGACATCAGGACTCAGAGTCCATATATCGGGCTCTCTT	8400
8341	CTCAACTGTCACTGAGAGGGAATCAGAACTCAGAGTCCATATACCAAGGCTCTCTT	8400
8401	GCCCGAGGAGGCCACATGTCATACATCTCGTAACTGAGAGACTTTTACGTGGAGGGCC	8460
8401	GCCCGAGGAGGCCACATGTCATACATCTCGCTGACTGAGAGACTTTTACGTGGAGGGCC	8460
8461	TATGTTCAACAGCAGGCGCAAACTCGGGGTACAGCGTTGCGGCCACAGCGGGTGCT	8520
8461	CATGTTTAAACAGCAGGGTTCAGCTCGGGGTACAGCGTTGCGGCCACAGCGGGTGCT	8520
8521	CACCACTAGCATGGGGAAACCATCACATGCTAGCTGAAAGCCCTTACGGCTTGTAAAGC	8580
8521	CACCTACTAGCATGGGGAAACCATCACATGCTAGCTGAAAGCCCTTACGGCTTGTAAAGC	8580
8581	TGACGGGATATTCGGGCCCAATGCTGGTATGGGGGATGACTTGGTTGTCTATCTAGA	8640

Copied from 09540843 on 05/19/2004

FMVSVSVSPDLPAIIERLHGLDAFSLHVTYPHELTRVSAFLKGLAPPLRAWS
RARAVASLIISHGAAVCRVYLFNVAVKLKLPLPEARLLDLSSWFTVGAGGDI
YLCVLRARPLLLLSLLLSVLGLLLPAR"

BASE COUNT 1940 a 2815 c 2709 g 2145 t 44 others

Query Match	86.2%;	Score	8367.8;	DB	14;	Length	9653;
Best Local Similarity	91.3%;	Pred. No. 0;					
Matches	8864;	Conservative	39;	Mismatches	750;	Indels	58;
Gaps	1;						
Qy	1	ACCCGCCCTTAATAGGGCGACACTCGGCATGAATCACTCCCTGTGAGGAATACTACTGT	60				
Db	1	ACCCGCCCTTAATAGGGCGACACTCGGCATGAATCACTCCCTGTGAGGAATACTACTGT	60				
Qy	61	CTTTCAGCAGAAAGCGTCTAGCCATGCGGTAGTATGAGTGTCTGATCAGCCTCCAGGCC	120				
Db	61	CTTTCAGCAGAAAGCGTCTAGCCATGCGGTAGTATGAGTGTCTGATCAGCCTCCAGGCC	120				
Qy	121	CCCCCTCCCGGAGAGCATAGTGGTCTCGGAAACCGGTGAGTACACCGGAATTGCGGG	180				
Db	121	CCCCCTCCCGGAGAGCATAGTGGTCTCGGAAACCGGTGAGTACACCGGAATTGCGGG	180				
Qy	181	AAGACTGGGTCTCTTTTGTGATAAACCACTATGCGCGGCCATTTGGGCGTGCCTCCG	240				
Db	181	AAGACTGGGTCTCTTTTGTGATAAACCACTATGCGCGGCCATTTGGGCGTGCCTCCG	240				
Qy	241	CAAGACTGCTAGCGGAGTGGGTTCGGAAGCGCTTGTGACTGCTCCCTGATAGG	300				
Db	241	CAAGACTGCTAGCGGAGTGGGTTCGGAAGCGCTTGTGACTGCTCCCTGATAGG	300				
Qy	301	TGCTTGCAGTGGCCCGGAGGTCTCGTAGACCGGTGACCATGAGCACAATCCTAAACC	360				
Db	301	TGCTTGCAGTGGCCCGGAGGTCTCGTAGACCGGTGACCATGAGCACAATCCTAAACC	360				
Qy	361	TCAAGAAAAACCAAGAAACCAACCGTTCGCGCAACAGAGCTTAAAGTTTCGGGCGG	420				
Db	361	TCAAGAAAAACCAAGAAACCAACCGTTCGCGCAACAGAGCTTAAAGTTTCGGGCGG	420				
Qy	421	CGGCAAGTCTGGGCGAGTACTTTGTCGCGCGAGGCGCCAGGTGCGGTGCG	480				
Db	421	CGGCAAGTCTGGGCGAGTACTTTGTCGCGCGAGGCGCCAGGTGCGGTGCG	480				
Qy	481	CGGCAAGGAAGACTTCGAGCGGTCCAGCACAAGTGAAGCGCGCCATCCCTAA	540				
Db	481	CGGCAAGGAAGACTTCGAGCGGTCCAGCACAAGTGAAGCGCGCCATCCCTAA	540				
Qy	541	AGATCGGCGTCCACTGGGAAACCAAGATACCCCTGGGCGCCCTATACGG	600				
Db	541	AGATCGGCGTCCACTGGGAAACCAAGATACCCCTGGGCGCCCTATATAGG	600				
Qy	601	GAATGAGGAGTCTGGGCGAGTGGGTCTGTCGCGCGAGGTTCCCGTCCCTCTTG	660				
Db	601	GAATGAGGAGTCTGGGCGAGTGGGTCTGTCGCGCGAGGTTCCCGTCCCTCTTG	660				
Qy	661	GGGCCCCAATGACCCCGGATAGTTCGCGCAAGTGGGTGAAGTCAATGATACCCCTAAC	720				
Db	661	GGGCCCCAATGACCCCGGATAGTTCGCGCAAGTGGGTGAAGTCAATGATACCCCTAAC	720				
Qy	721	GTGCGGTCTTCCGACTCATAGGGTACATCCCTGTCTGGGCGCCCGCTCGGCGCGT	780				
Db	721	GTGCGGTCTTCCGACTCATAGGGTACATCCCTGTCTGGGCGCCCGCTCGGCGCGT	780				
Qy	781	CGCAGAGCTCTCGCGATGCGGTGAGTTCCTGAGGAGCGGGTTAAATTTGCAACAGG	840				
Db	781	CGCAGAGCTCTCGCGATGCGGTGAGTTCCTGAGGAGCGGGTTAAATTTGCAACAGG	840				
Qy	841	GAATCTACCGGTGCTCTTTTCTATCTCTTCTGCGCCCTGCTGCTCTGCTACCCAC	900				
Db	841	GAATCTACCGGTGCTCTTTTCTATCTCTTCTGCGCCCTGCTGCTCTGCTACCCAC	900				
Qy	901	CCCGGTCTCGGCTCGGAGTGAAGAACATCAGTACCGGCTACATGCTGACTAACGACTG	960				
Db	901	CCCGGTCTCGGCTCGGAGTGAAGAACATCAGTACCGGCTACATGCTGACTAACGACTG	960				

Db	8818		CAGAGACCCCTACCACTCCCAATCGCCCGGCTGCTGGAAACAGTCAGACACTTCCTCCTGT	8877
Qy	8881		CAATTCATGCTCGGAAACATCATCATGTAAGCCCGACCATATGGCTCGCATGCTCT	8940
Db	8878		CAATTCATGCTCGGAAACATCATCATGTAAGCCCGACCATATGGCTCGCATGCTCT	8937
Qy	8941		GATGACACACTTCTTCTCCATTCATTCATGCTCAAGACACGCTCGACAGCACTCAACT	9000
Db	8938		GATGACACACTTCTTCTCCATTCATTCATGCTCAAGACACGCTCGACAGCACTCAACT	8997
Qy	9001		TGAGATGTCAGGAGCGGTGATCCGCTGAGTCCCTGGACCTCCAGCTCAATTAAG	9060
Db	8998		TGAAATGTCAGGATCGGTGATCCGCTGAGTCCCTGGACCTCCAGCTCAATTAAG	9057
Qy	9061		GTTACATGCGCTTGAAGCTTTTCTTGACACATACATCTCCCAAGCTGACAGCGGT	9120
Db	9058		GTTACAGCGGCTTGAAGCTTTTCTTGACACATACATCTCCCAAGCTGACAGCGGT	9117
Qy	9121		GGCTTCAGCCCTCAGAAACCTTGGCGCGCCACCCCTCAGAGCGTGAAGAGCGGCGCAG	9180
Db	9118		GGCTTCAGCCCTCAGAAACCTTGGCGCGCCACCCCTCAGAGCGTGAAGAGCGGCGCAG	9177
Qy	9181		TGCAGTCAGGCGCTCCCTCATCTCCGCTGGGGGAGAGCGCGCTTTCGCGTCAATCT	9240
Db	9178		TGCAGTCAGGCGCTCCCTCATCTCCGCTGGGGGAGAGCGCGCTTTCGCGCGGTATCT	9237
Qy	9241		CTTCAATGGGCGGTGAAGCAAGCTCAAACTCACTCACTCCATCCCGGAGCGGCTCT	9300
Db	9238		CTTCAATGGGCGGTGAAGCAAGCTCAAACTCACTCACTCCATCCCGGAGCGGCTCT	9297
Qy	9301		GGATTTATCAGCTGGTTCACCGCTCGCGCGCGCGCGCGGCGACATTTATCAGAGCGTGC	9360
Db	9298		GGATTTATCAGCTGGTTCACCGCTCGCGCGCGCGCGCGGCGACATTTATCAGAGCGTGC	9357
Qy	9361		GGCTGCGGACCCGCTTATGCTCTTTCGCTTACTCTTCTTCTTCTTCTTCTTCTTCTT	9420
Db	9358		GGCTGCGGACCCGCTTATGCTCTTTCGCTTACTCTTCTTCTTCTTCTTCTTCTTCTT	9417
Qy	9421		TTTCTTACTCCCGCTCGGTAGAGCGGACACATTAAGCTTACTTCACTCACTGATGCTC	9480
Db	9418		TTTCTTACTCCCGCTCGGTAGAGCGGACACATTAAGCTTACTTCACTCACTGATGCTC	9477
Qy	9481		CTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	9540
Db	9478		CTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	9537
Qy	9541		TTTTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	9600
Db	9538		TTTTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	9583
Qy	9601		TACTT	9660
Db	9584		T-CCTT	9642
Qy	9661		GAGCCGATGACGACAGAGTCCGCTTAACTGCTCTCTCTGACATCATGT	9711
Db	9643		GAGCCGATGACGACAGAGTCCGCTTAACTGCTCTCTCTGACATCATGT	9693

RESULT 12
 AF169004
 LOCUS
 DEFINITION
 Hepatitis C virus isolate G2aK3 polyprotein gene, complete cds.
 ACCESSION
 AF169004
 VERSION
 AF169004.1
 GI:6707283
 KEYWORDS
 Hepatitis C virus
 ORGANISM
 Hepatitis C virus
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 Hepacivirus.
 REFERENCE
 1 (bases 1 to 9653)
 Kurihara, C., Ishiyama, N., Nishiyama, Y., Fukushi, S., Kageyama, T.,

Katayama, K. and Miura, S.
 Molecular characterization of hepatitis C virus genotype 2a from
 the entire sequences of four isolates
 J. Med. Virol. 64 (4), 466-475 (2001)
 21361470
 11468731
 2 (bases 1 to 9653)
 Katayama, K. and Kurihara, C.
 Direct Submission
 Submitted (14-JUL-1999) Basic Research Division, BioMedical
 Laboratories, Inc., 1361-1 Matoba, Kawagoe, Saitama 350-1101, Japan
 Location/Qualifiers
 1. .9653
 /organism="Hepatitis C virus"
 /mol_type="genomic RNA"
 /isolate="G2aK3"
 /db_xref="taxon:11103"
 /note="genotype: 2a"
 341. .9442
 /codon_start=1
 /product="polyprotein"
 /protein_id="AAF25612.1"
 /db_xref="GI:6707284"
 /translation="MSTKPKPQKTKRNTNRRPODKFPGGGQIVGGVYLLPRRPRRL
 GVRATKTSERSQPRRQPIPKORRSTGKSGKPGYWPPLYGNEGGLGAGWLLSPRG
 SRPSWGTDPHRSRNKGVIDLTTCGADLMDGXI PVVGA PLGGVARALAHGVRVLED
 GVNATGNLPGCSFSIFLLALLSCITIPVSAQVNTSTGYMTNDCNSDITSCALY
 AVLVPGCVPCCKVMKSRWIPVSNVAVRPGALTQGLRTHIDMVMSATLSCALY
 VGDLCGGVMAAQMFIVSPHWFVQECNSIYPTITGHRMAMDMNMNSPATMIL
 AYAKRVPIVIDIIGAHVQWFGFLAYFSMOGAWKVVVILLAAAGVDAKTXAGXX
 AKTVSGIXILFXPGAKNGKQLINSNSWHINRLNCLNDSKTFGLASLFTYHSFNSS
 GCPRLSACRNLAEFRIGWTLQYENVTNEDMBPCYCHVPKCGKVPVPAESVCGPV
 YCFPSVWVGTDBLGVPTTWGENEDVLLASTRPPSSGSGFCWTWNSTGCTFKAC
 GAPPXRXDFNASTDLCDPDCFRKPEATIMKCGSGPWLTPCLDVPYPRMLHYPC
 TVNTYIFKIRYVGGVEHLTAACNFSRGYCNLDRDQGVSPILLSTFWAILLPC
 YTDLPALSTGLIHQNTVDQVMYGLSPALTKYVRWVWVLLFLILADARVACLW
 MLILQAAALEKLVLHAAASACNGFLVFIIFIAAYIKGRXVPLATVSLTGLW
 PFCLLALPOAVAYDGSVHGOIGALLILITLTFTLTPAYKTLISRLCMLCYLLTL
 GEMVWAPSMQARGDGLIWAATIFCPGVVFDITKMLLAVLGRGLYLLRDALTRVP
 YFVRAHLRMTAVRHLAGGRIYQMALALGRWTGTIYIDHLLTPMSWAASGLRDLA
 VAVPEITFSPMEKVTWGAETAAACDILHGLPVSARGRLILGAPDGYTSKRWLL
 APIYAAQOTRGLLGAIVVSMTRDKTEQAGEIOVLSTQTSFGLTSXGLVLYTHG
 AGNTLAGSGVPTQMSAEGDLVGPSPPTKSLPCTCGADVLYLVTRNADVIPA
 RRGDKRGALLSPRLSTLKGSGGSPVLCPRGHAVGI FRAACVSRGVAKSIDIFVEA
 LAIVTRSTFSDNSTPPAVPQTYQVGLHAPTGSKSTKVPVAAQGYKVLVNPVS
 LAITGFGAYLSRAHGINIRGTVTGAAITSTYTGKELADGGCGGADYIILCD
 ECHSVDSITILIGITVLDQAEATVRLTVLATATPPGVTTPHPNIEVLAQGEIIP
 FYGKAIPLSYIKGGRHLIFCHSKKKDELAALRGMLNAVAYRGLDVSIVTQGLV
 VVATDALMTGTGDFDSVIDCNVAVTQVDFSLDPTFTITQTPQDASRSORRG
 TGRGLGIYRVYSTGERASGMFDSVLCCEYDAGAAVVELTPVETVRLRAVFNTPGL
 PVCODHLEFWEAVFTGLTHIDAHFSLQKSGENFAYLVQAQYTCARAKAPPSWDI
 MWKLTURFLVGTPTLLYRLGSVNEVTHPVTKYIATCMQADLEVMSTVSLVLAG
 GVLAAVAYCLATGCVSILGRHXNQVAVAPDEKVEYDFEDECESRAALLBEGQ
 RIAMLSKIKIQLLQKASQADIQPAVQASRPVQEFWAKHMNFIISQIYLAGLST
 LPGNPVAMVMAFSAALTSPLSTTILNLLGWLASQIAPACATGFWVSGVGA
 VGSIGLKVILILAGYAGISGALVAFKIMSGEKPSMEDVVNLLPGLISGALVGV
 ICAAILRRHVGPGEQAVQWNRLLAFASRGNHVAPTHVVTESDASQVTLGLSITIT
 SLRLRLHNWITTEDCIPGAGSWLRDMDVWVCSILIDFKNWLTSKLFPKMPGLPLISQ
 KGYGVWAGTGMITTRCPGANISGNVRLGMSITGPKTCHNTWHGTFPINCYSQ
 VPKAPNPKTAIWRVAASEYAEVTHGYSYITGLTDNLKVPCLPSPPEFSSWVDG
 QIHRFAPTPKFRDEVSFCVGLNSFVGSOLPCDPRDPTDLTSLMTPDPSHITAEAA
 ABRLARGSPSEASSASOLSAISRLATCTTCKTYDVMDVANDLPMGDVTRIESES
 KIVLSDLSWVEGSDHEPSIPSEYMLPRTKFPALPVPWARPVPVPSWKRDPY
 QPPTVAGCALPSPKETPTPPRRRTVGLSESTIGDALQQLAIKTFGPPSTSGXS
 TGAADASGRTFPDELTPSTSTSPMPLEGPDPDPLEPQVPELQVPEQGVKVP
 GSDSGSWSTCEEDSVCCSMYSWTGALITPCSPSEKLPINFXSNLSRLYHNKVP
 CTSKASLRKAKVTEDRMVLDHSDVLDIKLAASKVSARLLTVSEACQLTPHS
 AKSKYFGCAKEVRSISRAVNHISKVMDLLEDSTPIPTIMAKNEVECPVTKGGK
 KAARLIVPDLGVRCERKALYDVTKLPQAVMGASGYFSQAQVREFLKAWADKK
 DPMGFSYDTRCFDSTVTERDIRTESIYQACSLPEEARTAHSLTULRGDVLVWISSQ
 GQTGCGYRCRASGLVITSMGNTITCYKALAAKCAKAAVIAPTMLVCGDVLVWISSQ
 ABEERNLRVFTAMTRYSPAGSDPPREYDLELITSCSNVSVARDPQGVRRYYLTR
 DPTPIARAAMETVRHFPVNSWLGNI IQAYPTIWRVLMVTHFFSILMAQDITLDQNLN

7741	CTTGAGCAATCCCTCTGTTGCGATATCAACAAGGTGTACTGTATACCAACAAAGAGCGC	7800
7738	CTTGAGCAATCGCTGTGTCGGGTACCAACAAGGTGTACTGTATACATCAAAAGAGCGC	7797
7801	CTCACTAAGGGCTAAAAGGTAACTTTTGTATAGGATGCAAGTGTCTGACTCTCTACTACGA	7860
7798	CTCACTGGGGCTAAAAGGTAACTTTTGTATAGGATGCGAGTGTCTGACGGCTATTATGA	7857
7861	CTCAGTCTTAAGGACATTAACTAGTCGGGCTCCAAGTGCACCGCAAGGCTCCTCACCAT	7920
7858	TTCAGTCTTAAAGGACATCAAGCTAGCGGCTCCAAGGTAGCGCGAGGCTCCTCACCTT	7917
7921	GGAGGAGGCTTGCAGATTAACTCCACCCCATCTCTCGAGATCTAAATATGGTTTGGGGC	7980
7918	GGAGGAGGATGCCAAATTGACTCACCCCATTTCTCAAGATCTAAGTATGGAATTTGGGGC	7977
7981	TAAGAGTGTCCGAGCTTTGTCCGGAGGGCCGTTAAACCACTCAAGTCCGTGTGGAAGGA	8040
7978	CAAGAGGTTCGAGCTTTGTCCGGAGGGCGCTCAACCACTCAAGTCCGTGTGGAAGGA	8037
8041	CCTCCTGGAGGACTCAGAAAACCAAAATCCCAACCAATATATGGCCAAAATATGAGGTGT	8100
8038	CCTCCTGGAAGACTCAGAAAACCAAAATCCCTACGACCATCATGGCCAAAATATGAGGTGT	8097
8101	CTGCGTGGACCCCAACGAGGGGGCAGAAAGACGCTCGCCTTATCGTTTACCCGTGACT	8160
8098	CTGCGTGCAGCCCAACGAGGGGGTAAAGAAAGACGCTCGCCTTATCGTTTACCCGTGACT	8157
8161	CGCGTTCAGGCTCTCGAGAAAGATGGCCCTTTATGACATTAACAAAAAATACCCGTGACT	8220
8158	CGGTTCAGGCTCTCGAGAAAGATGGCCCTTTATGATGTCAACAAAGCTTCTCTCAGGC	8217
8221	GGTATGGGGCTTCTTATGATTCAGTATTTCCCGCTCAGCGGGTAGAGTTTCTCTT	8280
8218	GGTATGGGGCTTCTTATGCTTCCAGTACTCCCCCGCTCAACGGGTGGAGTTCTCTCTT	8277
8281	GAAGCATGGCGGGAAGAAAGACCTTATGGGTTTTTCGTATGATATCCGATGCTTTGA	8340
8278	GAAGCATGGCGGGAAGAAAGACCTTATGGGTTTTTCGTATGATATCCGATGCTTTGA	8337
8341	CTCAACCGTCTACTGAGAGAGACATCAGAGCTGAGAGTCCATATATCGGCGCTGCTCCTT	8400
8338	CTCAACCGTCTACTGAGAGAGACATYAGAACTTGAAGAGTCCATATATCCGGGCTGCTCCT	8397
8401	GCCCGAGGAGGCCACACTGCGCATACACTCGCTAACTGAGAGACTTTACGTGGGAGGGCC	8460
8398	GCCGAGGAGGCCACACTGCGCATACATTTGCTGACTGAGAGACTTTACGTGGGAGGGCC	8457
8461	TATGTTCAACAGCAAGGGCCAAAACCTGGGGTACAGCGGTTGCCCGGCCAGCGGGTGCT	8520
8458	CATGTTCAACAGCAAGGGCCAGACCTGGGGTACAGGGGTTGCCCGGCCAGCGGGTGCT	8517
8521	CACCACTAGCATGGGGAACCATCACATGCTACGTGAAAGCCTTTAGCGGCTTGTAAAGC	8580
8518	CACCACTAGCATGGGGAACCATCACATGCTACGTRAAGCCCTAGCGGCTTGCAGGC	8577
8581	TGCAAGGATAATCGCGCCCAAACTGCTGGTATGCGCGCATGACTTGGTTGTCATCTCAGA	8640
8578	TGCGGGATAGTTGCACCCCAAAATGCTGATATGCGCGCAGCACTTGGTGTCTCATCTCAGA	8637
8641	AAGCAGGGGACCGAGAGAGACGAGCGGAACTTGAGAGCCTTCAAGGAGGCTATGACCCAG	8700
8638	AAGCAGGGGACCTGAGGAGGACGAGCGGAACTTGAGAGCTTCAAGGAGGCTATGACCCAG	8697
8701	GTATTTCCCTCTCTGGTGTACCCCGCCAGACCGGAGTATGATCTGGAGCTGATACATC	8760
8698	GTATTTCCCTCTCTCGCGGACCCCGCCAGACCCGGAATATGACCTGGAGCTGATACATC	8757
8761	TTGCTCTCAAAATGTCTGTGGCGCTGGGCCCAACAGGCCCGCCAGATACTACCTGAC	8820
8758	TTGCTCTCAAAATGTCTGTGGCACGGGACCAACAGGGCCCGCCAGATACTACCTGAC	8817
8821	CAGAGACCTTACCACTCAATCGCCCGGGCTGCTGGGAAACAGTTTGAACATCCCTCTGT	8880

4438	DB		GGAGGTGGCCCTCGGCGAGGAGGCGAGATCCCTTTCTATGGAGGCGCAATTCCTCGTGC	4497
4501	QY		TTATCATCAAGGAGGAGACATCTGATCTTCTGCCATTCAGAAAGAGTGTGACGAGCT	4560
4498	DB		TTATCATCAAGGAGGAGGACATCTGATCTTCTGCCATTCAGAAAGAGTGTGACGAGCT	4557
4561	QY		CGCGGGCGCCCTTTCGGGGTATGGGCTTGAACCTCAGTGGCATACTACAGAGGGTGGACGT	4620
4558	DB		CGCGGGCGCCCTTTCGGGGCATGGCTTGAACCTCAGTGGCATACTACAGAGGGTGGACGT	4617
4621	QY		CTCCGTAATACCAACTCAGGAGGAGCATGATGGTGGCGCCAGCAGCCCTCATGACAGG	4680
4618	DB		CTCTACAATAACCAACTCAGGAGGAGCATGATGGTGGCGCCAGCAGCCCTCATGACGCG	4677
4681	QY		GTATATCTGGGACTTTGACTCCGTGATCGACTGCAAGCTAGCGGTCACTCAAGTTGTAGA	4740
4678	DB		GTACACTGGAGACTTTGACTCAGTGTGACTGCAAGCTAGCGGTCACTCAAGTTGTAGA	4737
4741	QY		CTTCAGTTTACACCCACATTTACCATTAACCAACACAGATTTGCTCCCTCAAGACGCTGCTC	4800
4738	DB		CTTCAGTTTGGACCCACCTTTACCATTAACCAACACAGATTTGCTCCCTCAAGACGCTGCTC	4797
4801	QY		ACGTAGCCAGCGCGGGGTGCGACGGGTAGGGAAGACTGGGCATTTATAGGTATGTTTC	4860
4798	DB		ACGTAGCCAGCGCGGGGTGCGACGGGTAGGGAAGACTGGGCATTTATAGGTATGTTTC	4857
4861	QY		CACGTGTGAGGAGCGCTCAGGAATTTTGAACGTAGTGTCTGTGAGTGTCTACGACGC	4920
4858	DB		CACGTGTGAGGAGCGCTCAGGAATTTTGAACGTAGTGTCTGTGAGTGTCTACGACGC	4917
4921	QY		AGGGGGCGCATGTGTATGAGTCTACACCATCGAGAGCACCGTCAAGGCTCAGGCGGTATTT	4980
4918	DB		AGGGGGCGCATGTGTATGAGTCTACACCATCGAGAGCACCGTCAAGGCTCAGGCGGTATTT	4977
4981	QY		CAACACGCCCGTTTGTCTGTGTCGCAAGACCATCTTGAGTTTGGAGGCGATTTTCAC	5040
4978	DB		CAACACGCCCGTTTGTCTGTGTCGCAAGACCATCTTGAGTTTGGAGGCGGTTTTCAC	5037
5041	QY		CGGCTTACACATAGATGCCACTTCTTCCCAACCAAGCAATCGGGGAAATTT	5100
5038	DB		CGGCTTACACATAGATGCCACTTCTTCCCAACCAAGCAATCGGGGAAATTT	5097
5101	QY		CGCATACTTAAAGCTTACAGGCTACAGTGTGGCTAGGGCCAAAGCCCGCCCGCTC	5160
5098	DB		CGCATACTTAAAGCTTACAGGCTACAGTGTGGCTAGGGCCAAAGCCCGCCCGCTC	5157
5161	QY		CTGGGACGTATGGAAGTGTGACTCGACTCAAGCCACACTCTGTGGGCCCCACACC	5220
5158	DB		CTGGGATGTATGGAAGTGTGACTCGACTCAAGCCACACTCTGTGGGCCCCACACC	5217
5221	QY		TCTCTGTACCGCTTGGGCTCTGTTACCAAGAGGTGACCTCACATCCCGTGGAGAA	5280
5218	DB		TCTCTGTACCGCTTGGGCTCTGTTACCAAGAGGTGACCTCACATCCCTGTGAGAA	5277
5281	QY		ATACATCGCACCTGATGACGAGCCGACTTGAAGTATGACGAGCAGACATGGGCTTTGGC	5340
5278	DB		ATACATCGCACCTGATGACGAGCCGACTTGAAGTATGACGAGCAGACATGGGCTTTGGC	5337
5341	QY		AGGGGGAGCTTGTGGCGCGCTCGCCGCTATGCTTGGGACCGGGTGTGTTTGCATCAT	5400
5338	DB		TGGGGGGCTTGTGGAGCGCTCGCCGCTATGCTTGGGAGCTGGGTGTGTTTCCATCAT	5397
5401	QY		CGGCGCTTGCATTAACAGAGCGGCTGTTGGCGGACCAAGAGGCTCTCTATGA	5460
5398	DB		CGGCGCTTGCATTAACAGAGCGGCTGTTGGCGGACCAAGAGGCTCTCTATGA	5457
5461	QY		GGCTTTTGTAGATGAGGAATGTGCTCTAGGCGGCTCTCATTTGAAGAGGGGAGCG	5520
5458	DB		GGCTTTTGTAGATGAGGAATGTGCTCTAGGCGGCTCTCATTTGAAGAGGGGAGCG	5517
5521	QY		GATAGCGGAGATGCTGAAGTCCAAGATCCAAGGCTTATTCAGCAAGCTTTCCAAAC	5580
5518	DB		GATAGCGGAGATGCTGAAGTCCAAGATCCAAGGCTTATTCAGCAAGCTTTCCAAAC	5577
5581	QY		TCAGACATACCAACCCACCTGTGAGGCTTTCATGGCCCAAGGTAGAACAAATTCCTGGGCCAA	5640
5578	DB		TCAGACATACCAACCCACCTGTGAGGCTTTCATGGCCCAAGGTAGAACAAATTCCTGGGCCAA	5637
5641	QY		ACACATGTGGAACCTTCATTTAGCGGCATCCAAATCTCTCGAGGACTATCAACACTGCGAGG	5700
5638	DB		ACACATGTGGAACCTTCATTTAGCGGCATCCAAATCTCTCGAGGACTATCAACACTGCGAGG	5697
5701	QY		GAACCTGCACTAGCTTCCATGATGGCGTTTCACTGCGCGCCCTCACAGCCCGTGTGCGAC	5760
5698	DB		GAACCTGCACTAGCTTCCATGATGGCGTTTCACTGCGCGCCCTCACAGCCCGTGTGCGAC	5757
5761	QY		AAGCACCACTATCTTCTCAACATTTTGGGGGCTGCTAGCATCCCAATTCACACACC	5820
5758	DB		TAGCACCACTATCTTCTCAACATTTTGGGGGCTGCTAGCATCCCAATTCACACACC	5817
5821	QY		CGCGGGGCGCACTGGCTTCGTTGTCACTGGGCTAGTGGGAGCTGCGGTAGGCAATATAGG	5880
5818	DB		CGCGGGGCGCACTGGCTTCGTTGTCACTGGGCTAGTGGGAGCTGCGGTAGGCAATATAGG	5877
5881	QY		CTTAGTAAAGTGTAGTGGACATCTTGGCAGGATATGCTGGGCGCATTTTCGGGGCTCT	5940
5878	DB		CTTAGTAAAGTGTAGTGGACATCTTGGCAGGATATGCTGGGCGCATTTTCGGGGCTCT	5937
5941	QY		CGTGGCATTCATGATCATGCTTGGGAGAGCCCTCCATGGAGGATGTCTCAACTTGTCT	6000
5938	DB		CGTGGCATTCATGATCATGCTTGGGAGAGCCCTCCATGGAGGATGTCTCAACTTGTCT	5997
6001	QY		GCCTGGAATTCCTCTCGGGTGCCTTGTAGTGGAGTCACTTGGCGGCGCATTTCTGGG	6060
5998	DB		GCCTGGAATTCCTCTCGGGTGCCTTGTAGTGGAGTCACTTGGCGGCGCATTTCTGGG	6057
6061	QY		CGACACGTGGGACCGGGGGAAGCGCGCTCAATGGATGAATAGACTCATTTGCCCTTTGC	6120
6058	DB		CGGCAACGTGGGACCGGGGGAAGCGCGCTCAATGGATGAATAGACTCATTTGCCCTTTGC	6117
6121	QY		TTCCAGAGGAATTCAGTCTCGCCCGCCACCACTACGTGAGCGAGTCCGATCGGCGAGCG	6180
6118	DB		TTCCAGAGGAATTCAGTCTCGCCCGCCACCACTACGTGAGCGAGTCCGATCGGCGAGCG	6177
6181	QY		TGTGACCCCAACTTCTTGGCTTCCCTTACCATTAACAGCTGCTCAGAGGACTCCCAACTG	6240
6178	DB		TGTGACCCCAACTTCTTGGCTTCCCTTACCATTAACAGCTGCTCAGAGGACTCCCAACTG	6237
6241	QY		GATTTACTAGGACTGCGCCCATTCCTGCGCGGCTCTGCTGCTCCGCGATGTGTGGGACTG	6300
6238	DB		GATTTACTAGGACTGCGCCCATTCCTGCGCGGCTCTGCTGCTCCGCGATGTGTGGGACTG	6297
6301	QY		GGTTTGCACCATCTTAACAGACTTTTAAATTTGGCTGACCTCCAAATTTATTCCTCAAGAT	6360
6298	DB		GGTTTGCACCATCTTAACAGACTTTTAAATTTGGCTGACCTCCAAATTTATTCCTCAAGAT	6357
6361	QY		GCCCGGCTTCCCTTCTTGTCTCTCTGTCAAAAGGGGTACAAAGGGGCTGTGGGCGGCACTGG	6420
6358	DB		GCCCGGCTTCCCTTCTTGTCTCTCTGTCAAAAGGGGTACAAAGGGGCTGTGGGCGGCACTGG	6417
6421	QY		CATCATGACACACAGGTGCTTGGCGGCGCAATATCTCTGGCAATGTCTGCTGGGCTC	6480
6418	DB		TATCATGACACACAGGTGCTTGGCGGCGCAATATCTCTGGCAATGTCTGCTGGGCTC	6477
6481	QY		CATGAGAAATCAGGGGCTTAAGACCTGATGAATATCTGCGAGGGGACCTTTTCTTATCAA	6540
6478	DB		TATGAGAAATCAGGGGCTTAAGACCTGATGAATATCTGCGAGGGGACCTTTTCTTATCAA	6537
6541	QY		TTGTATACGAGGAGGCTAGTGTGCGGAAACCCGCGGCAAACTTTAAGGTGCCATCTG	6600
6538	DB		CTGTTTACGAGGAGGCTAGTGTGCGGAAACCCAGCTTAACTTCAAGACCGGCACTCTG	6597
6601	QY		GAGGGTGGCGGCTCAGAGTACCGGAGGTGACGAGGCTGATACCACTACATACATAC	6660
6598	DB		GAGGGTGGCGGCTCAGAGTACCGGAGGTGACGAGGCTGATACCACTACATACATAC	6657

Qy	2281	ATGCAATTTCACTCGTGGGGATCGTTGCAACTTTGGAGGACAGAGACAGAAGTCAACTGTTC	2340
Db	2278	GTGTAATTTCACTCGCGGGATCGTTGCGACTTTGGAGGACAGGAGTCAACTGTTC	2337
Qy	2341	TCCCTTTGTTGACATCCACCAAGGAATGGCCATTTTACCTTGTCTTACTTCGGACCTGCC	2400
Db	2338	TCCCTTTGTTGACATCCACCAAGGAATGGCCATTTTGCCTGTCTATACTTCAGACCTGCC	2397
Qy	2401	CGCCTTGTGCACTGGTCTTTCTCCACCTCCACCACCAACATCGTGGAGCTACAATTTCAATGTA	2460
Db	2398	TGCTTTGTGCACTGGCCTTTCTCCACCTCCACCACCAACATTTGTGGAGCTACAATACATGTA	2457
Qy	2461	TGGCCTATCACTGCTCCCTCACAAAATACATGCTCCGATGGAGTGGGTAAATATCTTTATTT	2520
Db	2458	TGGYTTGTCAACGCGCRCTCAAAAATACGTCGTCGCGATGGAGTGGGTAGTCTCTTATTT	2517
Qy	2521	CCTGCTCTTAGGGACCGCAGGGTTTGGCGCTGCTTATGATGCTCATCTTTGTTGGGCA	2580
Db	2518	CCTGCTCTTAGCGACCGCAGGGTCTGGCGTGGCTGTGATGCTCATCTTTGTTGGGCA	2577
Qy	2581	GGCCGAAGCAGCACTAGAGAAGCTGGTTCATCTTCACGCTGCGAGCGCAGCTAGCTGCAA	2640
Db	2578	AGCCGAGCAGCACTRGAGAAGCTGGTGGTCTCTGCAAGCTGCGAGCGCGCAGCTGCAA	2637
Qy	2641	TGGCTTCCTATATTTTGTGTCATCTTTTTCGTGGCTGCTTGTGTACATCAAGGGTCCGGTAGT	2700
Db	2638	TGGCTTCCTATATTTTGTGTCATCTTTTTCGTGGCTGCTTGGCAVATCAAGGGTCCGGCGGT	2697
Qy	2701	CCCCTTAGCTACTTATCCCTCACTGGCGCTGTGCTCTTTAGCCTACTGCTCTTAGCATTT	2760
Db	2698	CCCCTTAGCTGCTTATCCCTTACTTGGCCTATGGTCTCTTCTGCTACTGCTCTCTAATACT	2757
Qy	2761	GGCCCAACAGGCTTATGCTATATGCAATCTGTGCAATGGCAGATAGAGCGGCTCTGTCT	2820
Db	2758	GGCTCAACAGGCTTATGCTATATGATGATCGTGCACGGGCAGATAGGCGTGGTTTGT	2817
Qy	2821	GGTAATGATCACTCTCTTTACTCTCACCCCGGGTATAAGACCCCTTCTCAGCCGGTTTTT	2880
Db	2818	GACATGATACCTTCTTTACACTCACCCCGGGATATAAGACCCCTTCTCAGCCGGTGTCT	2877
Qy	2881	GTGGTGGTTGTCTATCTTTGACCCCTGGGGGAAGCTATGGTCCAGAGTGGGCAACACC	2940
Db	2878	GTGGTGGTTGTCTATTTTCTGACCCCTGGGGGAAGCATTGGTCCAGGAATGGGCAACATC	2937
Qy	2941	TATGAGTGGCGGTGGCGGTGATGGCATCATATGGCCGCTGGCCATATTTCTACCCAGG	3000
Db	2938	CATGACGGCGGTGGCGCGCGATGGCATCATATGGCCGCGCCACCATATTTCTGCCCGGG	2997
Qy	3001	TGTGGTGTGTGACATAAACCAAGTGGCTTGTGGCGGTGCTTGGGCGCTCTTACCTCCCTAAA	3060
Db	2998	CTTRGTGTGTGACATAAACCAAGTGGCTTGTGGCGGTGCTTGGGCGCTTACCTCCCTAAG	3057
Qy	3061	AGGTGCTTTGACGCGCGTGGCTGATCTTGGTCAAGGGCTCACGCTCTACTGAGAGATGGCAC	3120
Db	3058	AGATGCTTTGACGCGCGTGGCCATATTTCTGTCAGAGCTCACGCTCTGTGAGRGTTGGCAC	3117
Qy	3121	CATGCAAGGCATCTCGGGGGGGCAGTACGTCAGATGGCGCTACTAGCCCTTGGCAG	3180
Db	3118	TATGGTAGGCATCTCGGGGGGGTAAAGTACGTCAGATGGCGCTGTTAGCACTTGGTAG	3177
Qy	3181	GTGGACTGGCACTTACATCTATGACCACTCACCCCTATGTTCGGAATTTGGGCTGTAGTGG	3240
Db	3178	GTGGACCGGCACCTACATTTATGACCACCTCACCCCTATGTTCGGAATTTGGGCTGTAGCGG	3237
Qy	3241	CCTCGGGACCTGGCGGTGGCGTTGAGCCTATCATCTTTCAGTCCGATGGAGAGAAGT	3300
Db	3238	CCTCGGGACTTTGGCGGTGGCTGTGGAAACCATCTTTCAGCCCGCATGGAGAAGAAGT	3297
Qy	3301	CATTGCTGGGAGCGGACAGCTGCTTGTGGGACATTTTACACCGCATTTCCCGTGTC	3360
Db	3298	CATCGCTGGGGGGCAGARTCGGCCCATGCGGGGACATTTATGCACTCCCGCTGTC	3357

Db 6841 CCTGAGCCCGACGACGATATTGAGCTCCATCTCAACAGATCCGCTCCCATATCACGGC 6900
Qy 6901 GGAGACTGACGCGCGGTTTACGCGGGGGTTCACCCCATCCGAGGCAAGCTCTCTCGC 6960
Db 6901 GGAGGCTGACGCGCGGCTTACGCGGGGATCACCCCGTCCGAAGCAAGCTCTCTCGC 6960
Qy 6961 GAGCCAGCTATCGGCACCACTCGCTGCGAGCCACTGCACCAACCCACCGCAAGSCCTATGA 7020
Db 6961 GAGCCAGCTATCGGCACCACTCGCTGCGAGCCACTGCACCAACCCACCGCAAGSCCTATGA 7020
Qy 7021 TGTGGACATGGTGGATGCTAACTGTTTATGGGGGGGATGTGACTCGATAGAGTCTGG 7080
Db 7021 TGTGGACATGGTGGATGCTAACTGTTTATGGGGGGGATGTGACTCGATAGAGTCTGA 7080
Qy 7081 GTCCAAAGTGTCTGTTCTGACTCTCTCGACCAATGCTCGAAGAGAGGAGCGACCTTGA 7140
Db 7081 GTCCAAAGTGTCTGTTCTGACTCTCTCGACCAATGCTCGAAGAGAGGAGCGACCTTGA 7140
Qy 7141 GCCTTCGATACCATCAGATATGCTTCCCAAGAGAGGTTCCCAACAGCTTTTACCGGC 7200
Db 7141 GCCTTCGATACCATCAGATATGCTTCCCAAGAGAGGTTCCCAACAGCTTTTACCGGC 7200
Qy 7201 CTGGGACGCGCTGATTAACCCACCGCTTGTGGAATCGTGGAAAAGGCGAGATTAACA 7260
Db 7201 CTGGGACGCGCTGATTAACCCCGCTTGTGGAATCATGGAAGAGGCGAGATTAACA 7260
Qy 7261 ACCGGCAGCTGTTGGGGGTGCTCTCCCTCTCTAGGAAACCCCGACGCTCCCC 7320
Db 7261 ACCGGCAGCTGTTGGGGGTGCTCTCCCTCTCTAGGAAACCCCGACGCTCCCC 7320
Qy 7321 AAGGAGCGCGGACAGTGGGCTTAAGTAGGAGTCCATAGGAGATGCCCTTCAACAGCT 7380
Db 7321 AAGGAGCGTGGACAGTGGCTTAACGAGAACACATAGGAGATGCCCTTCAACAGCT 7380
Qy 7381 GGCCATTAAGTCTTTGGGAGCCCCCCCCAGCGGCGATTCAGGCTTTTCAACGGGGC 7440
Db 7381 GGCCATTAAGTCTTTGGGAGCCCCCCCCAGCGGCGATTCAGGCTTTTCAACGGGGC 7440
Qy 7441 GGGCGTGGCGATTCGGGAGTCCAGAGCTCTCTGATGAGTGGCCCTTTCCGAGACAGG 7500
Db 7441 GGGCGTGGCGATTCGGGAGTCCAGAGCTCTCTGATGAGTGGCCCTTTCCGAGACAGG 7500
Qy 7501 TTCCATCTCTTCATGCCCTTCGAGGGGAGCTTTGGAGATCCAGACTCGAGCCTGA 7560
Db 7501 TTCTATCTCTCTCCATGCCCTTCGAGGGGAGCCCGGGACCAAGCTTAGAGCCTGA 7560
Qy 7561 GCAGGTAGACCCCAACCCCGGCGGGGGTGGAGCTCCCGGCTCGGACTCGGG 7620
Db 7561 GCAGGTAGAGCTTCAGCTCCCGGCGGGGAGGAGTGTCTCTGGCTCGGACTCGGG 7620
Qy 7621 GTCTGGTCTACTTGTCTCGAGGAGGACGCTCCGCTGCTGCTGCTCCATCTCATCTC 7680
Db 7621 GTCTGGTCTACTTGTCTCGAGGAGGATGACTCCGCTGCTGCTGCTGCTCATCTATTC 7680
Qy 7681 CTGGACCGGGCTCTAATACTCTTGTAGTCCGAGAGGAGAGTTCAGGATTAACCC 7740
Db 7681 CTGGACCGGGCTCTAATACTCTTGTAGTCCGAGAGGAGAGTTCAGGATTAACCC 7740
Qy 7741 CTTGAGCACTCCCTGTTGGATATCAACAAGGTGTACTGTACCAACAAGAGCGC 7800
Db 7741 CTTGAGTAACTCGCTATTCGATATCAACAAGGTGTACTGTACTACATAAAGAGCGC 7800
Qy 7801 CTCACCTAAGGGCTAAAGGTAATTTGATAGGATGCAAGTGTCTGCTCTCTACTACGA 7860
Db 7801 CTCATTGAGGGCCAGAGGTAATTTGACAGGATGCAAGTGTCTGACGCCCATTTATGA 7860
Qy 7861 CTCAGTCTTAAAGGACATTAAGTACGCGCTTCCAGGTTCACCGAAGGCTCTCACCAT 7920
Db 7861 CTCAGTCTTAAAGGACATTAAGTACGCGCTTCCAGGTTCAGGCAAGGCTCTCACCTT 7920
Qy 7921 GGAGGAGGCTTGCAGTTAAACCCCAACCTTCTGCAAGATCTAAATATGGGTTGGGGC 7980
Db 7921 GGAGGAGGCTTGCAGTTAAACCTTCTGCAAGATCTAAATATGGGTTGGGGC 7980

Qy 7981 TAAGGAGTTCGACGCTTCTCGGGAGGSCGTTAAACACATCAAGTCCGTGTGGAAGA 8040
Db 7981 TAAGGAGTTCGACGCTTCTCGGGAGGSCGTTAAACACATCAAGTCCGTGTGGAAGA 8040
Qy 8041 CCTCTGGAGGACTCAGAAACCAATTCCTCAACACCAATTTATGGCCAAATAATGAGGTGT 8100
Db 8041 CCTCTGGAGGACTCAGAAACCAATTCCTCAACACCAATTTATGGCCAAATAATGAGGTGT 8100
Qy 8101 CTGGCTGGACCCCAACAAGGGGGGCAAGAACAGCTGCTGCTTATGCTTTACCTGACCT 8160
Db 8101 CTGGCTGGACCCCAACAAGGGGGGCAAGAACAGCTGCTGCTTATGCTTACCTGACCT 8160
Qy 8161 CGGCGTCAAGGCTCGGAGAGTGCCTTTATGACATTAACAAAATCTTCTCAGGC 8220
Db 8161 CGGCGTCAAGGCTCGGAGAGTGCCTTTATGATGTACAAAAGCTTCTCAGGC 8220
Qy 8221 GGTGATGGGGCTTCTTATGGATTCCAGTATTTCCCGCTCAGCGGTAGAGTTCCTCTT 8280
Db 8221 GGTGATGGGGCTTCTTATGGATTCCAGTATTTCCCGCTCAGCGGTAGAGTTCCTCTT 8280
Qy 8281 GAAAGATGGGGCGGAAAAGAGGACCTTATGGGTTTTCGTATGATACCCGATGCTTGA 8340
Db 8281 AAAGGCTAGGGCGGAAAAGAGGACCTTATGGGTTTTCGTATGATACCCGATGCTTGA 8340
Qy 8341 CTCACCGTCACTGAGAGAGACATCAGGACTGAGGAGTCCATATATCGGGCTGCTCTT 8400
Db 8341 TTTAAACCGTCACTGAGAGAGACATTAGAATCTAGAGTCCATATACCGGGCTGCTCTT 8400
Qy 8401 GCCGAGGAGGCCACACTGCGCATACACTGCTAACTGAGAGACTTTACGTGGGAGGCC 8460
Db 8401 GCCGAGGAGGCCACACTGCGCATACACTGCTAACTGAGAGACTTTACGTGGGAGGCC 8460
Qy 8461 TATGTTTCAACAGCAAGGGCCAAACCTGCGGTTTACAGGCTTCCCGCGCAGCGGGTGT 8520
Db 8461 CATGCTCAACAGCAAGGGTTCAGACTGCGGATACAGGGGTTTCCCGCGCAGCGGGTGT 8520
Qy 8521 CACCACTAGCATGGGAAACCATCAGATGCTGCTGAAAGCTTTAGCGGCTTTGAAGC 8580
Db 8521 CACCACTAGCATGGGAAACCATCAGATGCTGCTGAAAGCTTTAGCGGCTTTGAAGC 8580
Qy 8581 TGCGGGATAATTCGCGCCCAAACTGCTGATGCTGCGGATGCTTGGTGTCTCATCTCAGA 8640
Db 8581 TGCGGGATAATTCGCGCCCAAACTGCTGATGCTGCGGATGCTTGGTGTCTCATCTCAGA 8640
Qy 8641 AAGCAGGGGACCTGAGAGGAGCGAGCGAAACCTGAGAGCTTTTACCGAGGCTATGACCA 8700
Db 8641 AAGCAGGGGACCTGAGAGGAGCGAGCGAAACCTGAGAGCTTTTACCGAGGCTATGACCA 8700
Qy 8701 GTATTCTGCCCTCTGCTGAGCCCGCCAGACCGGAGTATGATCTGGAGCTGATTAACATC 8760
Db 8701 GTATTCTGCCCTCTGCTGAGCCCGCCAGACCGGAGTATGATCTGGAGCTGATTAACATC 8760
Qy 8761 TTGTTCTCTCAATGTTCTGTGGCGCTGGGCGCCCAAGCGCCCGCAGATACTACTGAC 8820
Db 8761 TTGTTCTCTCAATGTTCTGTGGCGCTGGGCGCCCAAGCGCCCGCAGATACTACTGAC 8820
Qy 8821 CAGAGACCTTACCTCAATGCTGCGCGGGTGTCTGGGAAACAGTTAGACACTTCCCTGT 8880
Db 8821 CAGAGACCTTACCTCAATGCTGCGCGGGTGTCTGGGAAACAGTTAGACACTTCCCTGT 8880
Qy 8881 CAATTTATGCTGGGAAACATCATCAGTACGCGCCGACCATATGGGCTCGCATGCTCT 8940
Db 8881 CAATTTATGCTGGGAAACATCATCAGTACGCGCCGACCATATGGGCTCGCATGCTCT 8940
Qy 8941 GATGACACATTTCTCTCAATTTCTCATGGCTCAAGACACGCTGGAGACCACTCAACTT 9000
Db 8941 GATGACACATTTCTCTCAATTTCTCATGGCTCAAGACACGCTGGAGACCACTCAACTT 9000
Qy 9001 TGAGATGTAAGGAGGCTGTACTCTGCTGAGTCTTGGAGCTTCCAGCTATTAATTGAAG 9060
Db 9001 TGAGATGTAAGGAGGCTGTACTCTGCTGAGTCTTGGAGCTTCCAGCTATTAATTGAAG 9060

Qy	4681	GTATACCTGGGACCTTTGACCTCCGTGATCGACGCAACGATAGCGGTCACTCAAGTTGTAGA	4740	Db	5761	TAGCACCACTATCTCTTCAACATATTTGGAGGCTGGCTGGGCTCCAAATCGCGCGCC	5820
Db	4681	GTATACCTGGGACCTTTGACCTCCGTGATCGACGCAACGATAGCGGTCACTCAAGTTGTAGA	4740	Qy	5821	CGCGGGGCGACCTGGCTTCTGTCAGTGGCCCTAGTGGAGCTGCGGTAGGCGAGTATAGG	5880
Qy	4741	CTTCAGTTTATAGACCCCACTTACACATPAAACACACAGATTGCTCCCTCAAGACGCTGTCTC	4800	Db	5821	CGCGGGGCGACCTGGCTTCTGTCAGTGGCCCTAGTGGGGGCTGCGGTAGGCGAGCATAGG	5880
Db	4741	CTTCAGCTTATAGACCCCACTTACATPAAACACACAGACTGTCCCGCAAGACGCTGTCTC	4800	Qy	5881	CTTAGTAAAGTGTAGTGGACATCTCTGCAGGGTATGCTGGGGGCAATTCGGGGGCTCT	5940
Qy	4801	ACGTAGCCAGCGCGGGGTGCGACGGGTAGGGAGACTGGGCAATTTATAGGTATGTTTC	4860	Db	5881	CTTGGTAAAGTGTAGTGGATATCTCTGCAGGGTATGCTGGGGCAATTCGGGGGCTCT	5940
Db	4801	ACGTAGCCAGCGCGGGGGCGACAGGTAGAGAACTGGGCAATTTATAGGTATGTTTC	4860	Qy	5941	CGTGCATTTCAAGATCATGCTTGGCGAAGCCCTCCATGAGAGATGTCGTCAACTGCT	6000
Qy	4861	CACGTGTAGCGAGCCCTCAGGAATGTTTGACAGTGTAGTCTCTGTGAGTGTACACGC	4920	Db	5941	CGTAGCATTTCAAGATCATGCTTGGCGAAGCCCTCCATGAGAGACGTCATCAACCTGCT	6000
Db	4861	CACGTGTAGCGAGCCCTCAGGAATGTTTGACAGTGTAGTCTCTGTGAGTGTACACGC	4920	Qy	6001	GCTCGAATTTCTGCTCCGGGTGCTTGTGTAGTGGAGTCACTCTGCGCGCCCATCTCTGCG	6060
Qy	4921	AGGGCCGCAATGATAGCTCACACATCGGAGACCAACGTCAGGCTCAGGGGCTATTTT	4980	Db	6001	GCTTGGGATCTTGTCTCCGGGGGCTTGTGTAGTGGAGTCACTCTGCGCGCCCATCTCTGCG	6060
Db	4921	AGGGGCTGCTTGTGTATGAGCTCACACATCGGAGACTACCGTCAAGGCTTAGAGCGTATTT	4980	Qy	6061	CGCACAGTGGGACCGGGGGAAGCGCCGTCCAAATGGATGAATAGATCTCATTTGCCCTTTC	6120
Qy	4981	CAACAGCCCGGCTTGTGCTGTGCAAGACCACTTGTAGTGTGAGGAGGAGTGTTCAC	5040	Db	6061	CGGCCAGTGGGCGCGGGGAGGGCGGCTCCAAATGGATGAATAGGCTCATTTGCCCTTTC	6120
Db	4981	CAACAGCCCGGCTTGTGCTGTGCAAGATCACTTGTAGTGTGAGGAGGAGTGTTCAC	5040	Qy	6121	TTCCAGAGGAATCAGTCTGCGCCCGCCACCTACGTGACGGAGTCCGATCGCTCGCAGCG	6180
Qy	5041	CGGCTTCACACATAGATCCCACTTCTTCCAAACAAAGCAATCGGGGGAATTTT	5100	Db	6121	TTCCAGAGGAACACAGTCTGCGCCCGCCACCTACGTGACGGAGTCCGATCGCTCGCAGCG	6180
Db	5041	TGGCTTCACACATAGATCCCACTTCTTCCAAACAAAGCAATCGGGGGAATTTT	5100	Qy	6181	TGTGACCCCAACTACTTGGCTTCCCTTACCATTAACAGCCTGCTCAGAGAGACTCCACAACTG	6240
Qy	5101	CGCATCTTAACAGCCTACAGGCTACAGTGTGGCTTAGGGCCAAAGCCCGCCCGCTC	5160	Db	6181	CGTGAACCAACTGCTTGGCTTCCCTTACTATTAACAGCCTGCTCAGAGAGACTCCACAACTG	6240
Db	5101	CGCATCTTAAGTACCTATCAGGCTACAGTGTGGCGCAGAGCCCAAGCCCGCTC	5160	Qy	6241	GATTACTGAGGAGCTGCCCATCCCATCGCGCGCTCGTGGCTCCCGCATGCTGTTGGGACTG	6300
Qy	5161	CTGGAGCTCATGTGGAAGTGTGCTGCTCAAGCCACACCTGCTGGGGGCGCCACACC	5220	Db	6241	GATTACTGAGAGCTGCGCCCATCCCATGCGCGGCTCATGGCTCCCGAGGTGTTGGGACTG	6300
Db	5161	CTGGAGCTCATGTGGAAGTGTGCTGCTCAAGCTCAAGCCACACCTGCTGGGGGCGCCACACC	5220	Qy	6301	GGTTTGCAACATCTTAAACAGACTTTTAAACACTGGCTGACCTCCAAATGTTTCCCAAGAT	6360
Qy	5221	TCTCTCTACCGCTTGGGCTCTGTTTACCAAGAGGTCAACCTCACAATCCCGTGACGAA	5280	Db	6301	GGTTTGCAACATCTTAAACAGACTTTTAAACACTGGCTGACCTCCAAATGTTTCCCAAGAT	6360
Db	5221	TCTCTCTACCGCTTGGGCTCTGTTTACCAAGAGGTCAACCTCACAATCCCGTGACGAA	5280	Qy	6361	GCCCGGCTCCCTTGTCTCTGTCACAAAGGGGTACAAAGGGGTGTTGGGCGCGGACTGG	6420
Qy	5281	ATACATCGCACTGCAATGAAGCCGACCTTGTAGGTATGACACAGCACATGCGGTCTTGGC	5340	Db	6361	GCCCGGCTCCCTTGTCTCTGTCACAAAGGGGTACAAAGGGGTGTTGGGCGCGGACTGG	6420
Db	5281	ATACATCGCACTGCAATGAAGCTGACCTTGTAGGTATGACACAGCACATGCGGTCTTGGC	5340	Qy	6421	CATCATGACACACAGCTGCTTTGGCGCGCAATATCTCTGGCAATGTCOGCTTGGGCTC	6480
Qy	5341	AGGGGAGTCTTGGCGCGCTGCGCGGTATGCTCGCGACCGGGTGTGTTGTCATCAT	5400	Db	6421	CATCATGACACACAGCTGCTTTGGCGCGCAATATCTCTGGCAATGTCOGCTTGGGCTC	6480
Db	5341	TGGGGAGTCTTGGCGCGCTGCGCGGTATGCTCGCGACCGGGTGTGTTTCCATCAT	5400	Qy	6481	CATGAGAACTACAGGGCTTAAACCTTGCATGAACTTTGGCAGGGGACCTTTTCCCATCAA	6540
Qy	5401	CGGCGGCTTGCACATTAACAGCGAGCGCTGCTTGGCGCGGACCAAGAGGTCTCTATGA	5460	Db	6481	TATGAGAACTTACAGGGCTTAAACCTTGCATGAACTTTGGCAGGGGACCTTTTCCCATCAA	6540
Db	5401	TGGCGGCTTGCACATTAACAGCGAGCGCTGCTTGGCGCGGACCAAGAGGTCTCTATGA	5460	Qy	6541	TTGTTACAGGGGCGGCTGCTGCGGCAAAACCGCGCGCAAACTTTTAAAGGTGCCCATCTG	6600
Qy	5461	GGCTTTTATGAGATGAGGAATGTGCTCTAGGGCGGCTCTCATTTGAAGGGGCGAGCG	5520	Db	6541	TTGTTACAGGGGCGGCTGCTGCGGCAAAACCGCGCGCAAACTTTTAAAGGTGCCCATCTG	6600
Db	5461	GGCTTTTATGAGATGAGGAATGTGCTCTAGGGCGGCTCTCATTTGAAGGGGCGAGCG	5520	Qy	6601	GAGGGTGGGGCTCAGAGTACCGGGAGGTGACGACGACGGGTCTATACCATACATAAAC	6660
Qy	5521	GATAGCCGAGTGTGAGTCCAGATCCAGGCTTATGCGCAAGCTTCCCAACCAAGC	5580	Db	6601	GAGGGTGGGGCTCAGAGTACCGGGAGGTGACGACGACGGGTCTATACCATACATAAAC	6660
Db	5521	AATAGCTGAGATGCTGAGTCCAGATCCAGGCTTATGCGCAAGCTTCCCAACCAAGC	5580	Qy	6661	AGGACTGACTACTGACAGTGTGAAAGTTCCTGTCAATTAACCTTCTCCAGAGTCTTTTTC	6720
Qy	5581	TCAAGACATACAAACCCACTGTGACGGCTTCAATGGCCCAAGGTAGAACTTTCTGGGCAA	5640	Db	6661	AGGACTGACTACTGACAGTGTGAAAGTTCCTGTCAATTAACCTTCTCCAGAGTCTTTTTC	6720
Db	5581	CCAGGATATACAAACCCCTGTGACGGCTTCAATGGCCCAAGGTAGAACTTTCTGGGCAA	5640	Qy	6721	CTGGGTGAGCGGAGTGCAGATCCATAGGTTTGGCCCGCACCAAGCCGTTTTTTCGGGA	6780
Qy	5641	ACAATGTGGAATCTTATAGCGGATCCCAATATCTCGAGGACTATCAACATGCGCAGG	5700	Db	6721	CTGGGTGAGCGGAGTGCAGATCCATAGGTTTGGCCCGCACCAAGCCGTTTTTTCGGGA	6780
Db	5641	ACAATGTGGAATCTTATAGCGGATCCCAATATCTCGAGGACTATCAACATGCGCAGG	5700	Qy	6781	TGAGGTCTCGTCTGCTGGGCTTAAATTCATTTGCTGCTGGGTCCCGAGCTTCTTTCGGA	6840
Qy	5701	GAACCTTCAGTGTGCTGAGTGGCTTCAAGTCCAGCTTCAACAGTCCGCTGTCAC	5760	Db	6781	TGAGGTCTCGTCTGCTGGGCTTCAATTCGTTGCTGCTGGGTCTCAGCTTCCCTGTGA	6840
Db	5701	GAACCTTCAGTGTGCTGAGTGGCTTCAAGTCCAGCTTCAACAGTCCGCTGTCAC	5760	Qy	6841	CCCTGAAACCCGACACAGAGCTATTTGATGTCATGCTGCTAAACAGATCCATCTCATATCACGGC	6900
Qy	5761	AAGCACCACTATCTTCTCAACATTTTGGGGGGCTGGCTAGCATCCCAAAATGACCAACC	5820				

Db 2461 TGGCTGTACCTGCTCTCAAAATATACATCGTTGGTGGAGTGGGTAGTACTCTTATT 2520
Qy 2521 CTGCTCTTAGCGGACCGCAGGCTTTGCGCTCTTATGATGCTCATCTTTGTTGGGCCA 2580
Db 2521 CTGCTCTTAGCGGACCGCAGGCTTTGCGCTCTTATGATGCTCATCTTTGTTGGGCCA 2580
Qy 2581 GGCCTAGAGCACTAGAGAAGCTGCTCATCTTTGCGACGCTGCGAGCGCACTAGCTGCAA 2640
Db 2581 GGCCTAGAGCACTAGAGAAGCTGCTCATCTTTGCGACGCTGCGAGCGCACTAGCTGCAA 2640
Qy 2641 TGGCTTCTATATTTGCTCATCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700
Db 2641 TGGCTTCTATATTTGCTCATCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700
Qy 2701 CCCCTTAGCTACCTATTTCCCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760
Db 2701 CCCCTTAGCTACCTATTTCCCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760
Qy 2761 GCGCCCAACAGGCTTATGCTTATGACGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820
Db 2761 GCGCCCAACAGGCTTATGCTTATGACGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820
Qy 2821 GGTAAATGATCACTCTCTTACTCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880
Db 2821 GATACGATTAACCTCTTCACTTATCCCGGGGTATAGATCTTCTCAGCCGGTCTCT 2880
Qy 2881 GTGGTGTGTGCTATCTTCTGACCTTGGGGAGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 2940
Db 2881 GTGGTGTGTGCTATCTTCTGACTTGGGGAGGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 2940
Qy 2941 TATGCAAGTGGCGGTGGCGGTGATGCTATATGCGCGGTGCGCATATTTACCCAGG 3000
Db 2941 CATGCAAGGCGGTGGCGGTGATGCTATATGCGCGGTGCGCATATTTACCCAGG 3000
Qy 3001 TGTGGTGTGATCAATCAAGTGGCTTCTTGGCGGTGCTTGGCGGTGCTTACCTCTAAA 3060
Db 3001 TGTAGTATTTGATCAATCAAGTGGCTTCTTGGCGGTGCTTGGCGGTGCTTATCTCTGGG 3060
Qy 3061 AGGTGCTTTGACGCGGTGCGGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3120
Db 3061 AGATGCTTGAACGCGGTGCGGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3120
Qy 3121 CATGCAAGGCTATCTCGCGGGGAGGTACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3180
Db 3121 TATGGTGAAGGCTATCTCGCGGGGAGGTACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3180
Qy 3181 GTGGACTGGCACTTACATCATGACCACTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3240
Db 3181 GTGGACCGGCACTTACATCATGACCACTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3240
Qy 3241 CCTGCGGACCTGCGCGGTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300
Db 3241 CCTGCGGACCTGCGCGGTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300
Qy 3301 CATTTGCTGGGAGCGGAGACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3360
Db 3301 CATCTGCTGGGAGCGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3360
Qy 3361 CGCCGCACTGCTGGGAGGTGCTTCTTGGCCAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420
Db 3361 CGCCGCACTGCTGGGAGGTGCTTCTTGGCCAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420
Qy 3421 GAGTCTTCTGCGCCCTCACTGCTTACCGCCAGCAGACGCTGCTGCTGCTGCTGCTGCTGCTGCT 3480
Db 3421 GAGGCTTCTTGGCCCTCACTGCTTACCGCCAGCAGACGCTGCTGCTGCTGCTGCTGCTGCTGCT 3480
Qy 3481 AGTGGTGAATGACGCGGCGGACAGACAGCAAGGCTGGGAAATTCAGTCTCTGCT 3540
Db 3481 AGTGGTGAATGACGCGGCGGACAGACAGCAAGGCTGGGAAATTCAGTCTCTGCT 3540
Qy 3541 CACAGTCACTAGTCTCTTCTCGGAACTCATCTCTCGGGGTTTGTGGACTGTCTACCA 3600
Db 3541 CACAGTCACTAGTCTCTTCTCGGAACTCATCTCTCGGGGTTTGTGGACTGTCTATCA 3600

Qy 3601 TGGAGCTGGCAAAAGACTCTGCGCGGCTCAGCGGCTCGGTCACGAGATGCTACTCCAG 3660
Db 3601 TGGAGCTGGCAAAAGACTCTGCGCGGCTCAGCGGCTCGGTCACGAGATGCTACTCCAG 3660
Qy 3661 TGTGAGGGGGACTTATGATAGGCTGCGCCAGCCCTCTGGGACTAAATCTTTGAGGCGGTG 3720
Db 3661 CGCCGAGGGGACTTGTAGGCTGCGCCAGCCCTCTCTGGGACTAAATCTTTGAGGCGGTG 3720
Qy 3721 CACGCTGAGGCGGTGCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3780
Db 3721 CACATGCTGAGGCGGTGCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3780
Qy 3781 AAGACGCGGGGCAAAAGCGGAGCGCTACTCTCTCCGAGACCTCTTTCCACCTTTGAAGG 3840
Db 3781 AAGACGCGGGGCAAAAGCGGAGCGCTACTCTCTCCGAGACCTCTTTCCACCTTTGAAGG 3840
Qy 3841 GTCTCTCAGAGGCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3900
Db 3841 GTCTCTCAGAGGCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3900
Qy 3901 TGTGCTGCTCTCGGGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3960
Db 3901 TGTGCTGCTCTCGGGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3960
Qy 3961 CGTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4020
Db 3961 CGTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4020
Qy 4021 TCAGTCTGCGGTACTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4080
Db 4021 TCAGTCTGCGGTACTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4080
Qy 4081 ATATGCTGCTCAGGGGTAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4140
Db 4081 ATATGCTGCTCAGGGGTAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4140
Qy 4141 GTTTCGGGCGGTACTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4200
Db 4141 GTTTCGGGCGGTACTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4200
Qy 4201 GACTGTGACGACCGGGGCGGCTCATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4260
Db 4201 AACCGTGAAGCGGGGAGTCCATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4260
Qy 4261 GGGCTGTGCGGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4320
Db 4261 GGGCTGTGCGGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4320
Qy 4321 TACCACTGCTTGGCATCGGAACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4380
Db 4321 TACCACTGCTTGGCATCGGAACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4380
Qy 4381 AACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4440
Db 4381 AACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4440
Qy 4441 GAGGTGCT 4500
Db 4441 GAGGTGCT 4500
Qy 4501 TTACATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4560
Db 4501 TTACATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4560
Qy 4561 CGCGCGGCGCTTCTGCGGGGTATGCGGCTTGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4620
Db 4561 CCGCGGCGCTTCTGCGGGGTATGCGGCTTGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4620
Qy 4621 CCGCGGCGCTTCTGCGGGGTATGCGGCTTGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4680
Db 4621 CCGCGGCGCTTCTGCGGGGTATGCGGCTTGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4680

301 TGCTTGGAGTGCCTGGGAGGCTCTGAGACCGTGCACCATGAGCACAATCTTAAACC 360
 301 TGCTTGGAGTGCCTGGGAGGCTCTGAGACCGTGCACCATGAGCACAATCTTAAACC 360
 361 TCAAGAAACCAAAAGAAACCAACCGTGCACCAAGAGCTTAAAGTTTCGGGGCGG 420
 361 TCAAGAAACCAAAAGAAACCAACCGTGCACCAAGAGCTTAAAGTTTCGGGGCGG 420
 421 CGGCGAGATGTTGGGAGGATATCTTGTTCGCGCGCAGGAGGCGCCAGGTTGGGTGCG 480
 421 CGGCGAGATGTTGGGAGGATATCTTGTTCGCGCGCAGGAGGCGCCAGGTTGGGTGCG 480
 481 CGGCGAAGGAAGACTTCGAGCGGTCCTGAGCGACGCTGGAAGAGCGCCAGCCATCCCAA 540
 481 CGGCGAAGGAAGACTTCGAGCGGTCCTGAGCGACGCTGGAAGAGCGCCAGCCATCCCAA 540
 541 AGATCGGCGCTGCATGCGGAAATCTTGGGAAACCAAGATACCCCTGGGCGCCCTATACGG 600
 541 AGATCGGCGCTGCATGCGGAAATCTTGGGAAACCAAGATACCCCTGGGCGCCCTATACGG 600
 601 GAATGAGGAGTCTGGGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 601 GAATGAGGAGTCTGGGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 661 GGGCCCCAATGACCCCGGATAGGTCGCGCAAGCTGGGTAAAGTCAATGATACCCCTAAC 720
 661 GGGCCCCAATGACCCCGGATAGGTCGCGCAAGCTGGGTAAAGTCAATGATACCCCTAAC 720
 721 GTGGGCTTTGCGGACCTCATGGGATACATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 721 GTGGGCTTTGCGGACCTCATGGGATACATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 781 CGGCAGAGCTCTCGCGACGCGTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGT 840
 781 CGGCAGAGCTCTCGCGACGCGTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGT 840
 841 GAATTTACCGGTTGCTCTTTTCTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 841 GAATTTACCGGTTGCTCTTTTCTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 901 CCGGCTCTCGCTGCGAAGTGAAGATACATGATACCGGCTACATGCTGATACGACTG 960
 901 CCGGCTCTCGCTGCGAAGTGAAGATACATGATACCGGCTACATGCTGATACGACTG 960
 961 CACCAATGAGAGATACCTGGAGCTTCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
 961 CACCAATGAGAGATACCTGGAGCTTCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
 1021 CCGGCTGCGAAGTGGGAAATGATCTCAGTGTGCTGATACCGGCTCTCACCGAATGTGGC 1080
 1021 CCGGCTGCGAAGTGGGAAATGATCTCAGTGTGCTGATACCGGCTCTCACCGAAGTGGC 1080
 1081 CGTGAGGCGGCGCGGCGCTTACGAGGCTTCCGAGCGGATGATGATGATGATGATGATGAT 1140
 1081 CGTGAGGCGGCGCGGCGCTTACGAGGCTTCCGAGCGGATGATGATGATGATGATGATGAT 1140
 1141 GTCCGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
 1141 GTCCGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
 1201 AGCCCAATGTTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
 1201 AGCCCAATGTTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
 1261 CATCTACCTGGTACCATGATCTGAGACACCGGATGAGGATGAGGATGAGGATGAGGATGAGG 1320
 1261 CATCTACCTGGTACCATGATCTGAGACACCGGATGAGGATGAGGATGAGGATGAGGATGAGG 1320
 1321 GCGCAGCGCTACCATGATCTTGGGCTGAGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
 1321 GCGCAGCGCTACCATGATCTTGGGCTGAGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
 1381 CATTAGCGGGGCTCATTTGGGCGCTCATGTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGG 1440

1381 CATTAGCGGGGCTCATTTGGGCGCTCATGTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGG 1440
 1441 GTGGGCGAAGTCTGTTGCTCATCTTCTGTTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1500
 1441 ATGGGCGAAGTCTGTTGCTCATCTTCTGTTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1500
 1501 TGTGGGGGCTTGGCGCGCAGACACCGGGGCGCTTCCAGGCTTATTTGACATGAGGCGCC 1560
 1501 CATTTGGGGTAGGCTGCGCGCAGCAGCAGCGGCTTGGCGGCGCTTCTTCTTCTTGGCGC 1560
 1561 CAGGCGAAGAAATCTGAGCTGTTAAACCAATGCGAGCTGCGACATCAACCGGCGGCGGCTT 1620
 1561 TAAGCAGAACATCCAGCTCATTAACCAACCAATGCGAGTTGGCAGTATCAACCGCAGCTGCCCT 1620
 1621 GAATGCAATGACTCTTCCGATACCGGCTTATCGGCTCTCTGTTCTTACACCTTACCGCTT 1680
 1621 GAATGCAATGACTCTTCCGATACCGGCTTATCGGCTCTCTGTTCTTACACCTTACCGCTT 1680
 1681 CAACTGCTCAGGATGCTCCGAAACGCAATGTCGCGCTGCGCGAGTATCGAGGCGCTTCCGGGT 1740
 1681 CAACTGCTCAGGATGCTCCGAAACGCTTGTGCGCTGCGCGGCTTATCCAGGCGCTTCCGGAT 1740
 1741 GGGATGGGCGCTTTCGAATATGAGGATTAATGTCAACCAATCCAGAGGATATGAGACCTTA 1800
 1741 AGGGTGGGCGACCTTTCGATACGAGGATAACGCTCACCAACCCAGAGAGATATGAGACCTTA 1800
 1801 TTGCTGGCACTACCCACCAAGGCGAGTGTGCGGCTGCTCCGCGAAGCTTGTGTGGCGCC 1860
 1801 CTGCTGGCACTACCCACCAAGGCGAGTGTGCGGCTGCTCCGCGAAGCTTGTGTGGCGCC 1860
 1861 AGTGTACTGTTTACCCCCAGGCGAGTGTGAGTGGGCGACGCGAGCGCTTGGAGCGCC 1920
 1861 AGTGTACTGTTTACCCCCAGGCGAGTGTGAGTGGGCGACGCGAGCGCTTGGAGCGCC 1920
 1921 CACTTACAGTGGGGGAGATGAGACAGATGCTTCTTATTTGAAAGAGCTTGGAGCGCC 1980
 1921 TACTTACAGTGGGGGAGATGAGACAGATGCTTCTTATTTGAAAGAGCTTGGAGCGCC 1980
 1981 GCTGGGCTCATGTTTGGGCTGCGAGTGAATGAACTCTTGGCTTACCAAGAGCTTGGCG 2040
 1981 GGGGGGCTCATGTTTGGGCTGCGAGTGAATGAACTCTTGGCTTACCAAGAGCTTGGCG 2040
 2041 CGCACCACTGCGGCTAGAGCTGACTTTCAACGCGCAGCAGCGAGCTGTTGTGCGCCAC 2100
 2041 CGCACCGGCTGCGGCTAGAGCTGACTTTCAATGCGCAGCAGCTGTTGTGCGCCAC 2100
 2101 GGAAGTGTAGGAGGATCTGATACCACTTACCTTCAATGCGGCTTGGGCGCTTGGCT 2160
 2101 GGAAGTGTAGGAGGATCTGATACCACTTACCTTCAATGCGGCTTGGGCGCTTGGCT 2160
 2161 CAGCGCAAGTGTGCTGATACCGCTTACCGCTTACCGCTTACCGCTTACCGCTTACCGCTTAC 2220
 2161 CAGCGCAAGTGTGCTGATACCGCTTACCGCTTACCGCTTACCGCTTACCGCTTACCGCTTAC 2220
 2221 CTATACCATCTTCAAAATTAAGATGATGTTGGGAGGCTTGGAGCAGAGCTTCAAGCTG 2280
 2221 CTATACCATCTTCAAAATTAAGATGATGTTGGGAGGCTTGGAGCAGAGCTTCAAGCTG 2280
 2281 ATGCAATTTCACTGTTGGGATGCTTGGCAACTTGGAGGAGCAGAGCAGAGCTTCAAGCTG 2340
 2281 ATGCAATTTCACTGTTGGGATGCTTGGCAACTTGGAGGAGCAGAGCAGAGCTTCAAGCTG 2340
 2341 TCTGTTGTTGCTGCTTCCACCAAGGAAATGGGCAATTTTACCTTGTCTTCTTCTGCGACCTGCC 2400
 2341 TCTGTTGTTGCTGCTTCCACCAAGGAAATGGGCAATTTTGTCTTCTTCTTCTGCGACCTGCC 2400
 2401 CGGCTTGTGAGCTGCTTCTTCCACCTTCCACCAAGGAAATGGGCAATTTTGTCTTCTTCTTCTG 2460
 2401 CGGCTTGTGAGCTGCTTCTTCCACCTTCCACCAAGGAAATGGGCAATTTTGTCTTCTTCTTCTG 2460
 2461 TGGGCTTATACCTTGGGCTTCAAAATTAAGATGCTGCTGCGAGTGGGAGTGGGTAATCTTATT 2520

```

QY 9301 GGATTTATCAGCTGGTTCACCGTCGGCGCGCGGGGCGACATTTATCAGAGCGTGC 9360
|
|
|
Db 9301 GGATTTATCAGCTGGTTCACCGTCGGCGCGCGGGGCGACATTTATCAGAGCGTGC 9360
|
|
|
QY 9361 GCGTGGCCGACCCCGCTATTGCTCTTTGGCTTACTCTCTATTTTGTAGGGGTAGCCCT 9420
|
|
|
Db 9361 GCGTGGCCGACCCCGCTATTGCTCTTTGGCTTACTCTCTATTTTGTAGGGGTAGCCCT 9420
|
|
|
QY 9421 TTTCTTACTCCCGCTCGGTAGAGCGGCACACATTAGCTACATCTCATAGTAACTGTCC 9480
|
|
|
Db 9421 TTTCTTACTCCCGCTCGGTAGAGCGGCACACATTAGCTACATCTCATAGTAACTGTCC 9480
|
|
|
QY 9481 CTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 9540
|
|
|
Db 9481 CTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 9540
|
|
|
QY 9541 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 9600
|
|
|
Db 9541 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 9600
|
|
|
QY 9601 TACTTTCTTCTGGTGGCTCCATCTTAGCCCTAGTACAGGCTAGCTGTGAAGTCCGT 9660
|
|
|
Db 9601 TACTTTCTTCTGGTGGCTCCATCTTAGCCCTAGTACAGGCTAGCTGTGAAGTCCGT 9660
|
|
|
QY 9661 GAGCCGCATGACTGCAGAGTCCGCTAACTGCTCTCTGACGATCATGT 9711
|
|
|
Db 9609 GAGCCGCATGACTGCAGAGTCCGCTAACTGCTCTCTGACGATCATGT 9659
|
|
|
RESULT 10
AF169005
LOCUS Hepatitis C virus isolate ND59 polyprotein gene, complete cds.
DEFINITION AF169005
ACCESSION AF169005
VERSION AF169005.1 GI:6707285
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
VIRUSES; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE 1 (bases 1 to 9700)
AUTHORS Katayama, K., Ishiyama, N., Nishiyama, Y., Fukushi, S., Kageyama, T.,
Katayama, K. and Miura, S.
TITLE Molecular characterization of hepatitis C virus genotype 2a from
the entire sequences of four isolates
J. Med. Virol. 64 (4), 466-475 (2001)
21361470
11468731
REFERENCE 2 (bases 1 to 9700)
AUTHORS Katayama, K. and Kurihara, C.
TITLE Direct Submission
Submitted (14-JUL-1999) Basic Research Division, BioMedical
Laboratories, Inc., 1361-1 Matoba, Kawagoe, Saitama 350-1101, Japan
Location/Qualifiers
1. .9700
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/isolate="ND59"
/db_xref="taxon:11103"
/notes="genotype: 2a"
341. .9442
/codon_start=1
/product="polyprotein"
/protein_id="AAF25613.1"
/db_xref="GI:6707286"
CDS
GVRATKTSERSQPRROP1PKDRSTGKSGKPGYPMPLPGNEGLGWAGWLLSPRG
/translation="MSTNPKPQKTKRNTSRPQDVKPPGGQIVGGVYLLPRRGPRL
SRPQWPTDPRSRNVNKGVIDTLTCGFADLMGYIPVVGAPLGGVAVARALAHGVRLD
GVNATDLPFGCSFIFLLALLSCITVPVSAVQVKNISYVMTNDCSNDSTITWQLQA
AVLHVPCCVNCPLSRWCIPVSPNVAQVRGALTOGLRAHIDIMVNMWSPATMIL
VGDLCCGVMLAAQMFIVSPQHHFVQECNCSTPGAITGHRMAWMMWNSPTAMIL
AVAMRVPEVIDI1ISAHGWVFMFLAYFSMQGAWAKVWVVIILLTAGVDAHTRSIAGSV

```

```

AHATSLAGLFTSGAKQNIQLINTNGSWHINRTALNCNDLSLNTGTGFIASLFTYTRFNS
GCPERLSACRGIQAFRIQWGLTLRYEDNVNTPEDMRPYCWHYPPKCGCIQVSARSQGPV
YCFTPSPVVGTTGDRGLGYPTVTWGENETDVEILANSTRPGGSMFCTWMSNGFTKTC
GAPPCRIRADFNASMDLLCPTDCFRKHPDATTIKCGSGFWLTPRCLVDITPYRMLHYPC
TINYTFIKRIMYGVGVEHLTAACNFTGRDCPCNEDRDSQSLPDLHSTTEWAILLPCPS
YSDLPALSTGLLHQNITVDQYMYGSPALTKYIVRWVWVLLLLADLADVCACLW
MLILGQAALEKLVLAHSAASANGFLYFVIFFVAWYIKGRAVPLAAYSLGTLW
PCLILLALPOQAYADASVHGTQVALLILITLFTLTPGYKILLSRSILMWCYLLTL
GEAMTOEVPFMQARGRDGIIWAATMFCPGVVFDTIKWLLAMLPAYILLRDLTRVP
VFAHALRLCTVVRHLAGGRYVQALLALGLRWGTGYIYDHLSPMSDWAASGKRLDLA
VAVEPIIFSPMEKKIVWGAETACDVLHGLPVSARGLEILLGLPAGSGLKRL
APIFAYAQOTRGLLGAIVVSMTRDKTQAGEIQVLSVTQSTGSLTSGVSLWTVHG
AGNKTLAGSRGPVTQWYSAGEDLVSGPCTGWLSEPCCTGCAVLYLVTRNADVIPA
RRGDKRGALLSPRLSTLKGSSGPPVLCRGHAVGI FRAAVCSRGVAKSILVIFVEA
LDVTRSPFSDNSTPPAVPQTYQVGYLHAPPTGSKSKTKVPVAYAAQGYKVLVNPV
AATLFGAYLSKAGHNIRGVTITGTSYISYTGKFLADGCGAGADIIICD
ECHAVDATILGVTVDQAEATAGVRLTLATATPPGVTTPPHNIEVALGQEGEIP
FYGRAIPLSYIKGGRHLIFCHSKKDBLSAALRSMGLNAVAYRGLDVSIVPTQGDV
VVVADALMTGYTGDSDVIDCNVAVTQVDFSLDPTITITQVPOAVSRQRRGR
TGRGLGIYRVVSTGERASGMFDSVWLCYCEYDAGAAYVELTPSETTVRLRAYFNTPL
PVCQDLHFEVMEVFTGLTHIDAHFSLQTSQSGENFAYLVAYQATVCARAKAPPSWDV
MVKTLRUKPLVGPPLLYRLGPVINEVTLTHPVTKYIATICMQADLEVMSTWVLG
GVLAAYACVCLATGCVSIIGRLHVNRAVAPDEKVEYAEFDEMECEASRAALIBEGQ
RIAEMLKSKIQLLQASQADIQPAVQASQPKVQFQWAKHMNFISGIVLAGLST
LPNPADVAMMAFSAALTSPSTSTTLLILNIGLWLSQIAIAPAGATIGVSGLVGAA
VGSIGLKVLDILAGYGAGISGALVAFKIMSGEKPSMEDVINLLPGLISGALVGV
ICAATLREHVGPGEQVQWNRLLJAFASRGNHVAPTHVYVTSASQRTQVLLGSLTIT
SLLRKHNITTEDCFPCAGSWLRVMDVCTILDTFKNLTSLKPFPMGLPFLSCQ
KGYGVNAGTGMITPCPGANISGNRLSGMRTGPKTMCNTWGTTPFPCNYTGGQC
VPKAPNFATIRVAASEYAEVTHQGSYSYITGLTDSLKVCQLPSPFPFSWYDGV
QIHRPAPTKPFPFDEVSFVGLNSFVVGSQLPCDPEPDADVLTFMGLDPSHITAEAA
ARRLARGSPSEASSASQAPSARATCTTHGKAYDVMDVANDLFLMGDVPTRISES
KYVLDLSDPNAERSDLEPSIPSEYMLPRNFPALPAWARPDPNPLPVSERKRPDY
QPTVAGCALPPTKPTSPPRRRRTVGNLNTGDLQQLAKAFQPPDLSGSLG
TCADAADSGSRTPPEDESSISMPLEGEPPDLEPEQVELOPPQGEVWP
GSDGSWSTCEEDDSVCCMSYSWTGALITPCSPSEKLPINLSNLSLRYHNKY
CTTSKSLRAKVTDFRMQVLDADYDOKI LAASKSARLLTBEACOLTTPHS
ARSKYGFGAKVRSLSGRAVNHISKVMDLLEDSTQPIPTTMAKNEVFCVDPKAGGK
KAARLIVYDGLVRCEKALYDVTKLPQAVMASGYFQSPQAVRVELLKAWEKK
DPMGFSYDTRCFDSTVTERDIRTESIYRACSLPEEAHTAHSRLERYLVGPMUNSK
QTCGRCRACRAGVLTSMGNTTTCYIKALAAKAAAGIAPVMTLVCGLDVLWISSQ
TEEBERNLRAFTTEAMTRYSPAPPPRPEYDELTITSCSNVSVXAGQRRRYLTR
DPTPLSRAAETVHRSVPNSWNLGNIIOYAPTIVWRVMTLHFFHILMAODTDLNLA
FEMGXVYVSPLDLPAILERLHGLDASLHMYSPHELTTRVASLRKLGAPELRAWS
PARAVRASLISRGGBRAALCGRYLFWAVKTKLTLPEARLLDLSSTFTVAGCGGDI
YHSVSRARPRLLLSLLELLLVGVGLFLPLAR"
BASE COUNT 1948 a 2845 c 2712 g 2190 t
ORIGIN

```

```

Query Match 86.8%; Score 8425.8; DB 14; Length 9700;
Best Local Similarity 91.9%; Pred. No. 0;
Matches 8920; Conservative 4; Mismatches 776; Indels 11; Gaps 2;

QY 1 ACCCGCCCTTAATAGGGCGACACTCCGCCATGAATCACTCCCTGTGAGGAACACTACTGT 60
|
|
|
Db 1 ACCCGCCCTTAATAGGGCGACACTCCGCCATGAATCACTCCCTGTGAGGAACACTACTGT 60
|
|
|
QY 61 CTTACGAGAGAAAGCGTCTAGCCATGGCGTTAGTAGTAGTGTCTGTACAGCCTCCAGGCC 120
|
|
|
Db 61 CTTACGAGAGAAAGCGTCTAGCCATGGCGTTAGTAGTAGTGTCTGTACAGCCTCCAGGCC 120
|
|
|
QY 121 CCCCCCTCCGGAGAGACCATAGTGTCTGCGGAACCGGTGAGTACACCGGAATTCGCGG 180
|
|
|
Db 121 CCCCCCTCCGGAGAGACCATAGTGTCTGCGGAACCGGTGAGTACACCGGAATTCGCGG 180
|
|
|
QY 181 AAGACTGGGTCTCTTTCTTGGATAAACCCACTCTATGCCCGGCCATTTGGCGGTGCCCG 240
|
|
|
Db 181 AAGACTGGGTCTCTTTCTTGGATAAACCCACTCTATGCCCGGCCATTTGGCGGTGCCCG 240
|
|
|
QY 241 CAAGACTCTAGCCAGTAGCGTGGGTTCGAAAGGCGCTTGTGTACTGCTCTGATAGG 300
|
|
|
Db 241 CAAGACTCTAGCCAGTAGCGTGGGTTCGAAAGGCGCTTGTGTACTGCTCTGATAGG 300
|
|
|

```


8161	CGGCGTCAGGCTGCGGAGAGATGGCCCTTTATGATGTCTACACAAAACTCCCTCAGGC	8220
8221	GGTGATGGGGCTTTCTTATGGAATTCAGTATTTCCCCGCTCAGCGGGTAGAGTTTCTCTT	8280
	GGTAATGGGGCTTTCTTATGGAATTCAGTATTTCCCCGCTCAGCGGGTAGAGTTTCTCTT	8280
8281	GAAAGCATGGCGGGAAGAAAGAGACCTATCGGTTTTTTCGTATGATACCCGATGCTTTGA	8340
	GAAAGCATGGCGGGAAGAAAGACCTATCGGTTTTTTCGTATGATACCCGATGCTTTGA	8340
8341	CTCAACCGTCACTGAGAGAGACATCAGGATGAGGAGTCCATATATCGGGCTGCTCCCTT	8400
	CACAAACCGTCACTGAGAGAGACATCAGGATGAGGAGTCCATATACCGGGCTGTTCCCT	8400
8401	GCCGAGGAGGCCACACTGCCCCATACTGCGTAACTGAGAGACTTTACGTGGAGGGCC	8460
	GCCGAGGAGGCCACACTGCCCCATACTGCGTAACTGAGAGACTTTACGTGGAGGGCC	8460
8461	TATGTTCAACAGCAAGGGCCAAAACCTGCGGGTACAGGCGTTTGC CGCGCCAGCGGGTGCT	8520
	TATGTTTAAACAGCAAGGGCCAGTCCCTGCGGTACAGGCGTTTGC CGCGCCAGTGGGGTGCT	8520
8521	CACCACATAGCATGGGGAACACCATCACATGCTACGTGAAAGCCTTAGCGGCTTGTAAGC	8580
	CACCACATAGCATGGGGAACACCATCACATGCTACGCAAAAGCCCTAGCGGCTTCAGGC	8580
8581	TGCAAGGATAATCGCGCCACAAATGCTGGTATGCGCGGATGACTTGGTTGTCAATCAGA	8640
	TGCAGGATAGTTGCGCCCCAATGCTGTATGCGCGGACGACTTGGTTGTCAATCAGA	8640
8641	AAGCAGGGGACCGAGGAGGACGAGCGGAACTGTAGAGCCTTTCAGGAGGCTATGACCG	8700
	AAGCAGGGGACCGAGGAGGATGAGCGGAACTGTAGAGCCTTTCAGGAGGCTATGACCG	8700
8701	GTATTCGCCCCCTCTCTGTGTGACCCCCCAGACCGGAGTATGATCTGGAGCTGATAACATC	8760
	GTATTCGCCCCCTCTCTGTGTGACCCCCCAGACCGGAGTATGATCTGGAGCTGATAACATC	8760
8761	TTGTCCTCAAAATGTGTCGTGGCGCTGGGCCCCACAAGCGCGCCAGATACTACTGAC	8820
	TTGTTCTCAAAATGTGTCGTGGCGTTGGGCCCCACAAGCGCGCCAGATACTACTGAC	8820
8821	CAGAGACCTTACCATCAATGCGCCGGCTGCTTGGGAAACAGTTAGACACTCCCTCT	8880
	CAGAGACCTTACCATCAAAATGCGCCGGCTGCTTGGGAAACAGTTAGACACTCCCTCT	8880
8881	CAATTCATGGCTGGGAAACATCATCCAGTAGCGCCCCGACCATATGSGCTCGCATGTCCT	8940
	CAATTCATGGCTGGGAAACATCATCCAGTAGCGCCCCGACCATATGSGCTCGCATGTCCT	8940
8941	GATGACACACTTCTCTCCATTTCTCATGGCTCAAGACAGCTGGACAGCAACTCAACTT	9000
	GATGACACACTTCTCTCCATCTCTCATGGCCCCAAGACACTCTGGACCCAGAACTCAATTT	9000
9001	TGAGATGACGAGCGGTGTACTCCGTGAGTCCCTTGGACCTCCCGACGTATAATTGAAG	9060
	TGAGATGATGAGCGGTGTACTCCGTGAAGTCCCTTGGACCTCCCGACCCATTAATTGAAG	9060
9061	GTTACATGGGCTTGACGCTTTTTTCTGCAACATACACTCCCGACGACTGACACGGGT	9120
	GTTACATGGGCTTGACGCTTTTTCTCTCTGCAACATACTCTCCCGACGAACTGACGCGGT	9120
9121	GGCTTCAGCCCTCAGAAAACCTTGGGGCGCACCTCTCAGAGGCTGGAAGAGCGGGCAG	9180
	GGCTTCAGCCCTCAGAAAACCTTGGGGCGCACCTCTCAGAGGCTGGAAGAGCTGCGGCGG	9180
9181	TGCAGTCAGGGCGTCCCTCATCTCCCGTGGAGGGAGCGCGCGTTTGGCGTGGTATCT	9240
	TGCAGTCAGGGCGTCCCTCATCTCCCGTGGAGGGAGCGCGCGTTTGGCGTGGTATCT	9240
9241	CTTCAATTTGGGCGGTGAAGACAAGCTCAAACTCACTCAATTTGCCGAAGCGGCTCCT	9300
	CTTCAATTTGGGCGGTGAAGACAAGCTCAAACTCACTCAATTTGCCGAAGCGGCTCCT	9300

Copied from 09540843 on 05/19/2004

1621	GAATGCAATGACTCTTCTTGCAACCGGCTTTATTCGGCTCTCTGTGTTCTATACCCACAGCTT	1680	QY
1621	GAATGCAATGAAATCCCTCGAACACCGGCTTTCTTGGCGCCCTGTGTTCTACGCCAAACGGCTT	1680	Db
1681	CAACTCGTCAGSAGTCCCGAAGCATGTCGCGCTGCGCAGTAGTATCGAGGCGCTTCGCGGT	1740	QY
1681	CAACTCGTCAGATGCTCTGAAACGCTGTCCGCTGCGCTGAAATATCGAGGAGTCTCCGAT	1740	Db
1741	GGGATGGGGCGCTTTCGAATATGAGATAATGTACCAATCCAGAGGATATGAGACCCCTA	1800	QY
1741	AGGATGGGGCAGATTGCATATACGAAATATGTACCAATTCAGAGGATATGAGACCATTA	1800	Db
1801	TTGCTGGCACTACCCACCAAGCAGTGTGGCTGTCTCCGCGAAGACTGTGTGTGGGCC	1860	QY
1801	TTGCTGGTACTACCCACCAAGCAGTGTGGCATAGTCCCGCGAGGTCTGTATGTGGCCC	1860	Db
1861	AGTGTACTGTTTACCCCCCAGCCAGTAGTAGTGGGCACGACCGACAGATCGGAGTGCC	1920	QY
1861	AGTGTACTGTTTACCCCCCAGCCAGTAGTAGTGGGCACGACCGACAGATCGGAGTGCC	1920	Db
1921	CACTTACAGTGGGGGAGATGAGACAGATGTCTTCTATTGAACAGCACTGCACCAAC	1980	QY
1921	TACCTACAGTGGGAGAGATGAGACAGATGTCTTCTATTGAACAGCACTGCACCAAC	1980	Db
1981	GCTGGGTCATGGTTGCGCTGCACGCTGAAGAACTTTCTGCTACACCAAGACTTGGG	2040	QY
1981	GTCGGCTCGTGGTTTGGCTGCACGCTGATGAATCCACTGGTTTACCACCAAAACATCGG	2040	Db
2041	CGCACCCCTGCCGTACTAGAGCTGACTTCAAACGACAGCAGGACCTGTGTGCCCCAC	2100	QY
2041	TGCACCACTTTCGCGCATTTAGAGCTGACTTCAAACGACAGCAGGACCTGTGTGCCCCAC	2100	Db
2101	GGACTGTTTTAGGAAGCATCTGATACCACTTACTCTAAATCGGCTCTGGGCCCTGGCT	2160	QY
2101	GGACTGTTTTAGGAAGCATCTGACCGCGCTTACTTCAGATGTGGCTCTGGGCCCTGGCT	2160	Db
2161	CACGCCAAGTGCCCTGATPCGACTACCCCTACAGGCTCTGGCATTACCCTGCACAGTTAA	2220	QY
2161	CACGCCAAGATGTTTGGTCGATTACCCCTACAGGCTCTGGCATTACCCTGCACAGTTAA	2220	Db
2221	CTATACCATCTTCAAAATACGATGTATGTGGAGGGTTCAGACACAGGCTCACGGCTGC	2280	QY
2221	CTATACCATCTTCAAAATACGATGTATGTGGAGGGTTCAGACACAGGCTTCGGGCTGC	2280	Db
2281	ATGCAATTTTCACTCGTGGGATCGTTGCAACTTCGAGGACAGACAGAACTCAACTGTC	2340	QY
2281	ATGCAACTTCACTCGCGGGATCGTTGGCACTTCGAGGACAGACAGAGATCAACTGTC	2340	Db
2341	TCCTTTGTGCACTCCACACGGAATGGCCATTTTACTCTGCTTACTCGGACTGCG	2400	QY
2341	TCCTTTGTGCACTCCACACGGAATGGGCCATTTTACCCTGCATTTACTCAGACTGCG	2400	Db
2401	CGCCTTGTGACTGGTCTTCTCCACCTCCAACAAACATCGTGGACGCTACATCATGTA	2460	QY
2401	CGCTTGTGACCGGCTCTTCTCCACCTCCAACAAACATCGTGGATGTACAATACATGTA	2460	Db
2461	TGGCCTATACCTGCCCTCAAAAATACATCGTCCGATGGGAGTGGGTAACTCTTATT	2520	QY
2461	TGGCCTGTACCTGCTCTTACAAAATACATCGTTCGATGGGAGTGGGTGTACTCTTATT	2520	Db
2521	CCTGCTCTTAGCGACCGCGGTTTGGCCTGTCTTATGGATGTCTCATCTTGTGGGCA	2580	QY
2521	CCTGCTCTTAGCGACCGCGGTTTGGCCTGTGGTATGGATGTCTCATTTTGTAGGCA	2580	Db
2581	GGCCGAAGCAGCACTAGAGAACTGGTCACTTTCGACGCTCGAGCGCAGCTAGCTGCAA	2640	QY
2581	GGCCGAAGCAGCACTAGAGAACTAGTCGTCTTGCACGCCCGCAGCGCGGCGAGCTGCAA	2640	Db
2641	TGGCTTCTATATTTTGTTCATCTTTTCGTGGCTGTGGTACATCATAGGCTCGGTA	2700	QY
2641	TGGCTTCTATATTTTGTTCATCTTCTCGTGGCTGTGGTACATCAAGGCGCGGCGGT	2700	Db
2701	CCCCTTAGTACTATTTCCTTCACTGGGCCCTGTGGTCTCTTTTAGCTACTGTCTCCTAGCAT	2760	QY

```

341..9442
/codon_start=1
/product="polyprotein"
/protein_id="BAB32874.1"
/db_xref="GI:13122266"
/translation="MSTTPKPKRTKNTNRPODVKFPGGQIVGVGYLLPRGGRL
GVRATKRSRQRRKKRQVITDKTGFADLMGYIVPGAPLGAGARALAGWLVRLSPRG
SRPSWPNDRPHRSRNGVDTLTGCFADLMGYIVPGAPLGAGARALAGWLVRLSPRG
GVNFATGNTLPGCSITPLALLSCITTPVSGAGVRNTSSVYVANDCANDSITWQLDA
AVLHPCVPCXKVGDTSCWCIPIVSPNVAVRQCAPLTRLGTRLTITDMVNAISCSALY
VGLDGCVMLAQMFIIVSPORHVMFOBCNCIYPGHVSGHRMAWDMWNMSPTATMIL
AYVMRPEVITDIIISGAHWGMFLGAYESQOGAWAKVVILLILAAAGVDAGTAVGGST
AHTTKGASQIFSGAGNQIQLNTNGSWHINRTALNCNSLNTGTFLAALFYANRFSNS
CGPERLSACNIEDIRIGMGTLOYENNVNTESDMRPYCWYIPRQCGIVPAKRSVCGPV
YCTPSPVVYVDTDLRGVLTPTWGENETDFLLNTRPPSGSFWGTCWNNSGTGFKTC
GAPPCRIRADFNASADLCTDCFRKHDPAAAYFCRSGSGWLPRLCLVDVYPIVRLHWYPC
TNAVYTFIKRIMVGGVHEHLAAACNTRDGRCDLEDRLSOLSPHLLSTTTEWAILPCT
YSDLPALSTGLLHLHQNIYDVQWYGLSPALTKIVRWBHWVLLFLLADARVACVAM
MLLILQGEAADEKLVIHLAASASNGFIYFVIFVAAVYIKGRAVPWATVSLTGLW
SFCLLALLQQQAYAYDTSAHGQIGATLLILITFTTPAYKTLVSRELVMLCYLLTL
GEAMVQEWAPLPGVWRGRDGIYIAWAAYFCGVGVFDITKMLLALGSPYLILDALTRVLA
YFVRHALLRICTVMWRHLAGCRVYQVOMALLAGRWTCGYIVDHLSPMSDMAANGRLDA
VAVEPIIFSPMEKKVIVMGAEATAACQDILHGLPVSRILGREVLPADGVTSGKWSLL
APITYAQQTRGLLGAIVSMSTGRDKTEQAGEIQVLSTVQTSFLGTSIAGVLTVYVHG
AGNKTLAGSRGPVTQWYSSAEGDLVEMPSPPGTKSLEPCTCGHADVILVTRNADYIIPA
RRRGDGRKALLSPRLSTLSSGGSPVLCPRGHAVGI FRAAVGSRGVAKSIDIPVET
LDIVTSRPSDNSTPAPVOTYOVGYLHAPTGTSGSKSTKVPVAYAAQGVKVLVNLSPV

```

[illegible]

QY	7261	ACGGGCACTGTTGCGGGCTGTGCTCTCCCTCTCTCTAGGAAACCCCGACGCTCCCC	7320	8341	CTCAACCGCTCACTGAGAGAGACATCAGGAGCTGAGGAGTCCATATATCGGGCTGTCTCT	8400
DB	7261	ACGGGCACTGTTGCGGGCTGTGCTCTCCCTCTCTCTAGGAAACCCCGACGCTCCCC	7320	8341	CTCAACCGCTCACTGAGAGAGACATCAGGAGTCCATATATCGGGCTGTCTCT	8400
QY	7321	AAGGAGCGCGGACAGTGGGCTTAAGTAGAGTCCATAGAGAGTGGCTTCAACAGCT	7380	8401	GCCGAGGAGGCGGACATGCGCATACATCGCTAACTGAGAGACTTTAGTGGAGGGCC	8460
DB	7321	AAGGAGCGCGGACAGTGGGCTTAAGTAGAGTCCATAGAGAGTGGCTTCAACAGCT	7380	8401	GCCGAGGAGGCGGACATGCGCATACATCGCTAACTGAGAGACTTTAGTGGAGGGCC	8460
QY	7381	GGCCATTAACTCTTTGGCCAGCCCCCCCCCAAGGGGGATTAGGGCTTTCCACGGGGC	7440	8461	TATGTTTCAACAGCAAGGGCCAAACCTGCGGGTACAGGGCTTCCGCGCCAGCGGGTGT	8520
DB	7381	GGCCATTAACTCTTTGGCCAGCCCCCCCCCAAGGGGGATTAGGGCTTTCCACGGGGC	7440	8461	TATGTTTCAACAGCAAGGGCCAAACCTGCGGGTACAGGGCTTCCGCGCCAGCGGGTGT	8520
QY	7441	GGCGCTGCGGATTCGCGGAGTCAGAGCGCTCTGATGAGTGGCCCTTTTCGAGAGCG	7500	8521	CACCACTAGCATGGGGAACACCATCACATGCTACGTTGAAGCTTTAGCGGCTTTGAAGC	8580
DB	7441	GGAGGAGCGGATTCGCGGAGTCAGAGCGCTCTGATGAGTGGCCCTTTTCGAGAGCG	7500	8521	TACCACTAGCATGGGGAACACCATCACATGCTATGTGAAGCTTTAGCGGCTTTGAAGC	8580
QY	7501	TTCCATCTCTTCATGCCCCCTCGAGGGGAGCTTGGAGATCCAGACTGGAGCTTGA	7560	8581	TGCAAGGATTAATCGCGCCCAAACTGCTGGTATGCGCGCATGATCTTGGTTGTCTCAGA	8640
DB	7501	TTCCATCTCTTCATGCCCCCTCGAGGGGAGCTTGGAGATCCAGACTGGAGCTTGA	7560	8581	TGCAAGGATTAATCGCGCCCAAACTGCTGGTATGCGCGCATGATCTTGGTTGTCTCAGA	8640
QY	7561	GCAGGTAGAGCCCAACCCCCCCCCCAGGGGGGGTGGAGCTCCGGCTCCGACTCGGG	7620	8641	AAGCCAGGGGACCGAGGAGGACGAGCGGAACTGAGAGCTTTCAGAGGCTTATGACCA	8700
DB	7561	GCAGGTAGAGCTTCAACCTCCCCCCCCCAGGGGGGGTGGTAAACCCCGGCTCAGGCTCGGG	7620	8641	GHGCCAGGGGACCGAGGAGGACGAGCGGAACTGAGAGCTTTCAGAGGCTTATGACCA	8700
QY	7621	GTCTGTGCTTACTTGTCTCGAGGAGGACGACTCCGTCGTGTGCTGCTCCATGTCTACTC	7680	8701	GTATTTCTGCCCCCTCTGCTGAGACCCCGCCAGACCGGAGTATGATCTGGAGCTGATAACATC	8760
DB	7621	GTCTGTGCTTACTTGTCTCGAGGAGGACGACTCCGTCGTGTGCTGCTCCATGTCTACTC	7680	8701	GTATTTCTGCCCCCTCTGCTGAGACCCCGCCAGACCGGAGTATGATCTGGAGCTGATAACATC	8760
QY	7681	CTGACCGGGGCTCTAATACTCTTGTAGTCCGAGAGGAGGAGTACCGATTAAACCC	7740	8761	TTGCTCTCTCAATGTGTGTGGGCTGGGCCCCAAGGCCCGCCAGATACTACTGAC	8820
DB	7681	CTGACCGGGGCTCTAATACTCTTGTAGTCCGAGAGGAGGAGTACCGATTAAACCC	7740	8761	TTGCTCTCTCAATGTGTGTGGGCTGGGCCCCAAGGCCCGCCAGATACTACTGAC	8820
QY	7741	CTTGAGCACTCCCTGTGCGATATCACACAGGTGTACTGTACCAACAAGAGCGC	7800	8821	CAGAGACCTTACCTCCAACTCGCCCGGGCTGCTGGGAAACAGTTAGACACTCCCTCTG	8880
DB	7741	CTTGAGCACTCCCTGTGCGATATCACACAGGTGTACTGTACCAACAAGAGCGC	7800	8821	CAGAGACCTTACCTCCAACTCGCCCGGGCTGCTGGGAAACAGTTAGACACTCCCTCTG	8880
QY	7801	CTCACTAAGGGCTAAAGAGTAACTTTGATAGATGCAAGTCTCGACTCTCTACTACGA	7860	8881	CAATTCATGGCTGGGAAACATCATCAGTACGCCCCCGACCATATGGGCTCGCATGCTCT	8940
DB	7801	CTCACTAAGGGCTAAAGAGTAACTTTGATAGATGCAAGTCTCGACTCTCTACTACGA	7860	8881	CAATTCATGGCTGGGAAACATCATCAGTACGCCCCCGACCATATGGGCTCGCATGCTCT	8940
QY	7861	CTCAGTCTTAAGGACATTAAGCTAGCGGCTCCAAAGTCAACGAGGCTCCCTCAACAT	7920	8941	GATGACACATCTTCTCTCATCTCATGCTCAAGACGCTGGAACAGAACTCAACT	9000
DB	7861	CTCAGTCTTAAGGACATTAAGCTAGCGGCTCCAAAGTCAACGAGGCTCCCTCAACAT	7920	8941	GATGACACATCTTCTCTCATCTCATGCTCAAGACGCTGGAACAGAACTCAACT	9000
QY	7921	GGAGGAGCTTCCAGTTAAACCCCAACCTCTGCAAGTCTAAATATAGGTTGGGCG	7980	9001	TGAGATGTACGAGCGGTGTACTCGTGAGTCCCTTGGAGCTCCAGCTTATATTTGAAAG	9060
DB	7921	GGAGGAGCTTCCAGTTAAACCCCAACCTCTGCAAGTCTAAATATAGGTTGGGCG	7980	9001	TGAGATGTACGAGCGGTGTACTCGTGAGTCCCTTGGAGCTCCAGCTTATATTTGAAAG	9060
QY	7981	TAAGGAGGCTCCGAGCTTTGCGGAGGGCGGTTAACCATCAAGTCCGTGTGGAAGA	8040	9061	GTATCATGGGCTTGACGCTTTTCTCTGCACACATACACTCCCGACGAACTGACAGGCT	9120
DB	7981	TAAGGAGGCTCCGAGCTTTGCGGAGGGCGGTTAACCATCAAGTCCGTGTGGAAGA	8040	9061	GTATCATGGGCTTGACGCTTTTCTCTGCACACATACACTCCCGACGAACTGACAGGCT	9120
QY	8041	CTCTCTGGAGGACTGAAACACCAATTCACCAACCATTTATGGCCAAAATAGAGTGT	8100	9121	GGCTTCAGGCTTCAGAAAACCTTGGGCGCCACCCCTCAGAGCGTGGAGAGCGGGCAG	9180
DB	8041	CTCTCTGGAGGACTGAAACACCAATTCACCAACCATTTATGGCCAAAATAGAGTGT	8100	9121	GGCTTCAGGCTTCAGAAAACCTTGGGCGCCACCCCTCAGAGCGTGGAGAGCGGGCAG	9180
QY	8101	CTCGGTGAGCCCAACAGGGGGGCAAGAGAGCTCGCTTATCGTTTACCTGACCT	8160	9181	TGCACTGAGGCGCTCCCTCATCTCCGCTGGGGGAGAGCGGCTTTGCGGCGATATCT	9240
DB	8101	CTCGGTGAGCCCAACAGGGGGGCAAGAGAGCTCGCTTATCGTTTACCTGACCT	8160	9181	TGCACTGAGGCGCTCCCTCATCTCCGCTGGGGGAGAGCGGCTTTGCGGCGATATCT	9240
QY	8161	CGCGCTCAGGCTCTCGAGAGATGGCCCTTTATGACATTTACACAAAATCTCTCAGGC	8220	9241	CTTCAATTTGGGCGGTGAAGACCAAGCTCAAACTCACTCACTTTCGCGGAGCGGCTCT	9300
DB	8161	CGCGCTCAGGCTCTCGAGAGATGGCCCTTTATGACATTTACACAAAATCTCTCAGGC	8220	9241	CTTCAATTTGGGCGGTGAAGACCAAGCTCAAACTCACTCACTTTCGCGGAGCGGCTCT	9300
QY	8221	GGTGTAGGGGCTTCTTATGATTTCCAGTATTTCCCGCTCAGCGGTTAGTCTCTT	8280	9301	GGATTTATCCAGCTGGTTCAACCGTCCGCGCGGGGCGGACATTTATCAAGCGGTGC	9360
DB	8221	GGTGTAGGGGCTTCTTATGATTTCCAGTATTTCCCGCTCAGCGGTTAGTCTCTT	8280	9301	GGATTTATCCAGCTGGTTCAACCGTCCGCGCGGGGCGGACATTTATCAAGCGGTGC	9360
QY	8281	GAAGCATGGGCGGAAAGAGACCTTATGGGTTTTTGTATGATACCGGATGCTTTGA	8340	9361	GGTGTCCGACCCGCTTATTTACTCTTGGGCTTCTCTCTCTTTTGTAGGGGTAGGCT	9420
DB	8281	GAAGCATGGGCGGAAAGAGACCTTATGGGTTTTTGTATGATACCGGATGCTTTGA	8340	9361	GGTGTCCGACCCGCTTATTTACTCTTGGGCTTCTCTCTCTTTTGTAGGGGTAGGCT	9420
				9421	TTTCTCTACTCCCGCTCGGTAGAGCGGACACATTTAGCTTACACTCCATAGCTTAAGTCTC	9480

5041	CGSCCTCACACATAGATGCCACTTCTTTTCCAAAACAAAGCAATCGGGGAAATTT	5100	6121	TTCCAGAGGAAACACAGTCGCGCCCAACCCACTACGTGACGAGTCGGATCGGTGCGAGCG	6180
5101	CGCATACTTAACAGCTACAGGCTACAGTGTGGCTAGGCGCAAGACCCCGCCCGCTC	5160	6181	TGTGACCCAACTACTTGGCTCCCTTACCATTAACAGCTGCTCAGAGACTCCACAACCTG	6240
5101	CGCATACTTAACAGCTACAGGCTACAGTGTGGCTAGGCGCAAGACCCCGCCCGCTC	5160	6181	TGTGACCCAACTACTTGGCTCCCTTACCATTAACAGCTGCTCAGAGACTCCACAACCTG	6240
5161	CTGGGAGCTCATGTGGAAGTGTGACTCGACTCAAGCCACACTCGTGGGCGCCACACC	5220	6241	GATTACTGAGGAGTCCCGCATCCCATGCGCGGCTCTGTGGCTCCGCGATGTGTGGACTG	6300
5161	CTGGGAGCTCATGTGGAAGTGTGACTCGACTCAAGCCACACTCGTGGGCGCCACACC	5220	6241	GATTACTGAGGAGTCCCGCATCCCATGCGCGGCTCTGTGGCTCCGCGATGTGTGGACTG	6300
5221	TCTCTGTATCCGCTTGGGCTCTGTATCCAAACGAGGTCAACCTCATCGCATCTGTGAGCA	5280	6301	GGTTTGCACCATCTTAACAGACTTTTAAAACTGGCTGACCTCAAAATGTGTCCAAAGAT	6360
5221	TCTCTGTATCCGCTTGGGCTCTGTATCCAAACGAGGTCAACCTCATCGCATCTGTGAGCA	5280	6301	GGTTTGCACCATCTTAACAGACTTTTAAAACTGGCTGACCTCAAAATGTGTCCAAAGAT	6360
5281	ATACATGCGCACCTGATGCAAGCGGCTTGAAGTCAATGACAGACATGGGTCTTGGC	5340	6361	GCCCGGCTCCCTTGTCTCTGTCAAAAGGGGTACAAAGGGCTGTGGCGCGGCACTGG	6420
5281	ATACATGCGCACCTGATGCAAGCGGCTTGAAGTCAATGACAGACATGGGTCTTGGC	5340	6361	GCCCGGCTCCCTTGTCTCTGTCAAAAGGGGTACAAAGGGCTGTGGCGCGGCACTGG	6420
5341	AGGGGAGTCTTGGCGGCTCGCGGCTATTCCTGGCGGCTGTGTGTCATCAT	5400	6421	CATCATGACCAACACGCGTCTTGGCGGCGGCAATATCTTGGCAATGTCTGGCTTGGGCTC	6480
5341	AGGGGAGTCTTGGCGGCTCGCGGCTATTCCTGGCGGCTGTGTGTCATCAT	5400	6421	CATCATGACCAACACGCGTCTTGGCGGCGGCAATATCTTGGCAATGTCTGGCTTGGGCTC	6480
5401	CGGCGCTTGCAATTAACAGGAGCGCTGTGCGCGGCAAGAGAGTCTCTATGA	5460	6481	CATGAGATCACGGGCTTAAGCTGATGATATCTGCGAGGGGACCTTTTCCATCAA	6540
5401	CGGCGCTTGCAATTAACAGGAGCGCTGTGCGCGGCAAGAGAGTCTCTATGA	5460	6481	CATGAGATCACGGGCTTAAGCTGATGATATCTGCGAGGGGACCTTTTCCATCAA	6540
5461	GGCTTTGTAGATGAGAGGAATGTGCTCTAGGGGCTCTCATTTGAAGAGGGGAGCG	5520	6541	TTGTTACAGGAGGCGAGTGTGCGGCAAAACCCGCGCAAACTTTAAGGTGCGCATCTG	6600
5461	GGCTTTGTAGATGAGAGGAATGTGCTCTAGGGGCTCTCATTTGAAGAGGGGAGCG	5520	6541	TTGTTACAGGAGGCGAGTGTGCGGCAAAACCCGCGCAAACTTTAAGGTGCGCATCTG	6600
5521	GATAGCGAGATCTGAAGTCCAAGATCCAAAGCTTATTCAGCAAGCTTCCAAACAAGC	5580	6601	GAGGCTGCGGCTCAGAGTACGCGAGGTGACGAGCAAGGCTCATACCACTACATAAC	6660
5521	GATAGCGAGATCTGAAGTCCAAGATCCAAAGCTTATTCAGCAAGCTTCCAAACAAGC	5580	6601	GAGGCTGCGGCTCAGAGTACGCGAGGTGACGAGCAAGGCTCATACCACTACATAAC	6660
5581	TCAAGACATCAACCCGCTGTGAGGCTTCTTGCGCCCAAGGTAGACAAATCTTGGGCGAA	5640	6661	AGGACTTACCACTGATTAACCTTGAAGTTCCTTGGCAACTACCTTCCAGAGTCTTTTC	6720
5581	TCAAGACATCAACCCGCTGTGAGGCTTCTTGCGCCCAAGGTAGACAAATCTTGGGCGAA	5640	6661	AGGACTTACCACTGATTAACCTTGAAGTTCCTTGGCAACTACCTTCCAGAGTCTTTTC	6720
5641	ACATGAGTGAATCTTATAGCGGCTCAATACCTCGAGGACTCAACACTGCGAGG	5700	6721	CTGGGTGAGCGGAGTGCAGATCCATAGTTCGCGCCCAACCGAAGCGCTTTTCGCGGA	6780
5641	ACATGAGTGAATCTTATAGCGGCTCAATACCTCGAGGACTCAACACTGCGAGG	5700	6721	CTGGGTGAGCGGAGTGCAGATCCATAGTTCGCGCCCAACCGAAGCGCTTTTCGCGGA	6780
5701	GAACCTGAGTACTTCAATGAGGCTTCAAGTCCGCGCTCAACAGTCCGCTGTCAAC	5760	6781	TGAGGTCTCGTTCGCTTGGGCTTAATTAATTTGCTGCGGTCCAGCTTCTTGGCGGA	6840
5701	GAACCTGAGTACTTCAATGAGGCTTCAAGTCCGCGCTCAACAGTCCGCTGTCAAC	5760	6781	TGAGGTCTCGTTCGCTTGGGCTTAATTAATTTGCTGCGGTCCAGCTTCTTGGCGGA	6840
5761	AAGCACCATCTCTCAACATTTTGGGGGCTGCTAGCATCCCAATTTGCAACACC	5820	6841	CCCTGAACCGGACACAGAGTATGATGTCATGCTAAAGAGTCCATCTCATATCAGGC	6900
5761	AAGCACCATCTCTCAACATTTTGGGGGCTGCTAGCATCCCAATTTGCAACACC	5820	6841	CCCTGAACCGGACACAGAGTATGATGTCATGCTAAAGAGTCCATCTCATATCAGGC	6900
5821	CGCGGGGCTTACCGGCTTCTGCTGCTAGTGGGCTGCTGCTAGGAGCATAGG	5880	6901	GGAGCTGACGCGGCTTTAGCGCGGCTTCAACCCCACTCCGAGGCAAGCTCTCGGC	6960
5821	CGCGGGGCTTACCGGCTTCTGCTGCTAGTGGGCTGCTGCTAGGAGCATAGG	5880	6901	GGAGCTGACGCGGCTTTAGCGCGGCTTCAACCCCACTCCGAGGCAAGCTCTCGGC	6960
5881	CTTAGGTAAAGTCTAGTGACATCTTGGGAGGATGTGTCGCGGCTTTTCGCGGCTCT	5940	6961	GAGCCAGCTATCGGCAACCATCGCTGCGAGCCACTGCAACCCACGAGGCAAGGCTATGA	7020
5881	CTTAGGTAAAGTCTAGTGACATCTTGGGAGGATGTGTCGCGGCTTTTCGCGGCTCT	5940	6961	GAGCCAGCTATCGGCAACCATCGCTGCGAGCCACTGCAACCCACGAGGCAAGGCTATGA	7020
5941	CGTGGCATCAAGATCATGTCTGGCGAGAGCCCTCCATGAGGAGTCTGTCATCTGCT	6000	7021	TGTGGACATGCTGGATGCTTAACTGTTTCAATGCGGGGCGATGTCATCGGATAGTCTGG	7080
5941	CGTGGCATCAAGATCATGTCTGGCGAGAGCCCTCCATGAGGAGTCTGTCATCTGCT	6000	7021	TGTGGACATGCTGGATGCTTAACTGTTTCAATGCGGGGCGATGTCATCGGATAGTCTGG	7080
6001	GCCTGGAATCTGCTCTCGGGTCTGCTGCTAGTGGGAGTCACTGCGCGGCTTCTGCG	6060	7081	GTCCAAAGTGGTCTGCTGAGCTCTCTCGACCCCAATGCTGCGAAGAAAGAGGCACTTGA	7140
6001	GCCTGGAATCTGCTCTCGGGTCTGCTGCTAGTGGGAGTCACTGCGCGGCTTCTGCG	6060	7081	GTCCAAAGTGGTCTGCTGAGCTCTCTCGACCCCAATGCTGCGAAGAAAGAGGCACTTGA	7140
6061	CCGACGCTGGGACCGGGGAGGCGCGCTCAATGATGATGATGATGATGATGATGATG	6120	7141	GCCTTCGATACCATCAGAAATACATGCTCCCAAGAGAGGTTCCTCCAGCTTTACCGGC	7200
6061	CCGACGCTGGGACCGGGGAGGCGCGCTCAATGATGATGATGATGATGATGATGATG	6120	7141	GCCTTCGATACCATCAGAAATACATGCTCCCAAGAGAGGTTCCTCCAGCTTTACCGGC	7200
6121	TTCCAGAGGAAATCACGCTCGCGCCCAACCCACTACGTGACGAGTCCGATCGTGCAGCG	6180	7201	CTGGGACGCGCTGATTAACACCCACCGCTTGTGGAAATCGTGGAAAGGCGAGATTAACA	7260
6121	TTCCAGAGGAAATCACGCTCGCGCCCAACCCACTACGTGACGAGTCCGATCGTGCAGCG	6180	7201	CTGGGACGCGCTGATTAACACCCACCGCTTGTGGAAATCGTGGAAAGGCGAGATTAACA	7260

8461 CATGTTCAACAGCAGCGGCAGACCTCGCGGTACAGCGGTGCGCGCCAGCGGGTCT 8520
 8521 CACACCTAGCATGGGAAACACCATCATCTGCTGAAAGCCTTAGCGGCTTGTAAAGC 8580
 8521 TACACTAGCATGGGAAACACCATCATCTGCTGAAAGCCTTAGCGGCTTGTAAAGC 8580
 8581 TGCAGGATTAATCGCGCCCAACAATGCTGGTATGCGGCGATGACTTGTGCTCATCTCAGA 8640
 8581 TGCAGGATTAATCGCGCCCAACAATGCTGGTATGCGGCGATGACTTGTGCTCATCTCAGA 8640
 8641 AAGCCAGGACCGAGAGGAGCAGCGGAACCTTGAGAGCCTTCCAGAGGCTATGACAC 8700
 8641 GAGCCAGGACCGAGAGGAGCAGCGGAACCTTGAGAGCCTTCCAGAGGCTATGACAC 8700
 8701 GTATTTCGCCCTCTGCTGAGACCGCCCGGAGCGGAGTATGATCTGGAGCTATGATC 8760
 8701 GTATTTCGCCCTCTGCTGAGACCGCCCGGAGCGGAGTATGATCTGGAGCTATGATC 8760
 8761 TTGCTCTCAATGCTGCTGGCGCTGGGCGCCACAGCGCGCGAGATCTACCTGAC 8820
 8761 TTGCTCTCAATGCTGCTGGCGCTGGGCGCCACAGCGCGCGAGATCTACCTGAC 8820
 8821 CAGAGACCTTACCTCAATCGCCCGGCTGCTCGGAAACAGTTAGACACTTCCCTCT 8880
 8821 CAGAGACCTTACCTCAATCGCCCGGCTGCTCGGAAACAGTTAGACACTTCCCTCT 8880
 8881 CAATTATGCTCGGAAACATCATCAGTACCGCCCGGAGCCATATGCGCTCGATGCTCT 8940
 8881 CAATTATGCTCGGAAACATCATCAGTACCGCCCGGAGCCATATGCGCTCGATGCTCT 8940
 8941 GATGACACCTTCTCTCAATCTCATGCTCAAGACACGCTGAGCAGAACTTCAACT 9000
 8941 GATGACACCTTCTCTCAATCTCATGCTCAAGACACGCTGAGCAGAACTTCAACT 9000
 9001 TGAGATGTAACGAGCGGTGACTCCGAGTCCCTTGGACCTCCAGCTTAAATTGAAG 9060
 9001 TGAGATGTAACGAGCGGTGACTCCGAGTCCCTTGGACCTCCAGCTTAAATTGAAG 9060
 9061 GTTACATGGCTTGACGCTTTTCTCTGACACATACACTCCCGAGCTGACAGGCT 9120
 9061 GTTACATGGCTTGACGCTTTTCTCTGACACATACACTCCCGAGCTGACAGGCT 9120
 9121 GGTTCAGCCCTCAGAAACTTGGGCGCCACCCCTCAGAGCTGGAAGCGCGGCACG 9180
 9121 GGTTCAGCCCTCAGAAACTTGGGCGCCACCCCTCAGAGCTGGAAGCGCGGCACG 9180
 9181 TGCAATGAGGCGTCCCTCATCTCCGTTGGGAGAGCGCGCTTGGCGTCAATCT 9240
 9181 TGCAATGAGGCGTCCCTCATCTCCGTTGGGAGAGCGCGCTTGGCGTCAATCT 9240
 9241 CTTCAATGAGGCGTGAAGCAAGCTCAAACTCACTTCCATGTCGGAAGCGCGCTCT 9300
 9241 CTTCAATGAGGCGTGAAGCAAGCTCAAACTCACTTCCATGTCGGAAGCGCGCTCT 9300
 9301 GGATTTATCCAGCTGTTTCACTGCGCGCGCGCGGCGGAGCATTTATCAGAGCTGTC 9360
 9301 GGATTTATCCAGCTGTTTCACTGCGCGCGCGCGGCGGAGCATTTATCAGAGCTGTC 9360
 9361 GCGTCCCGACCCCGCTTATGCTTCTTGGCTTACTCTTCTTGTGGGCTAGGCT 9420
 9361 GCGTCCCGACCCCGCTTATGCTTCTTGGCTTACTCTTCTTGTGGGCTAGGCT 9420
 9421 TTTCTACTCCCGCTCGGTAGAGCGGACATTAAGCTACACTCCATGACTAACTGTC 9480
 9421 TTTCTACTCCCGCTCGGTAGAGCGGACATTAAGCTACACTCCATGACTAACTGTC 9480
 9481 CTTT 9540
 9481 CTTT 9540
 9541 TT 9588
 TT

9541 TT 9588

RESULT 8
 LOCUS I12861
 DEFINITION Sequence 2 from patent US 5428145.
 ACCESSION I12861
 VERSION I12861.1 GI:910242
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 9589)
 AUTHORS Okamoto, H. and Nakamura, T.
 TITLE Non-A, non-B, hepatitis virus genome, polynucleotides, polypeptides, antigen, antibody and detection systems
 JOURNAL Patent: US 5428145-A 2 27-JUN-1995;
 FEATURES
 Location/Qualifiers
 source 1..9589
 /organism="unknown"
 BASE COUNT 1968 a 2820 c 2635 g 2166 t
 ORIGIN

Query Match 92.4%; Score 8970.4; DB 6; Length 9589;
 Best Local Similarity 96.0%; Pred. No. 0;
 Matches 9202; Conservative 0; Mismatches 386; Indels 0; Gaps 0;

QY 1 ACCCGCCCTAATAGGGGCGACATCCCGCATGAATCACTCCCTGTGTAGGAACACTACTGT 60
 DB 1 ACCCGCCCTAATAGGGGCGACATCCCGCATGAATCACTCCCTGTGTAGGAACACTACTGT 60
 QY 61 CTTACGCGAAGAGCGTCTAGCCATGCGGTGTAGTAGTGTGTGTGTGTGTGTGTGTGTGT 120
 DB 61 CTTACGCGAAGAGCGTCTAGCCATGCGGTGTAGTAGTGTGTGTGTGTGTGTGTGTGT 120
 QY 121 CCCCCTCCCGGAGAGCCATAGTGTCTGCGAAACCGGTGAGTACACCGGAATTCGCGGG 180
 DB 121 CCCCCTCCCGGAGAGCCATAGTGTCTGCGAAACCGGTGAGTACACCGGAATTCGCGGG 180
 QY 181 AAGACTGGTCTCTTTCTTGGATAAACCCACTCTATGCGCGGTCAITTTGGCGGTGCGCCCG 240
 DB 181 AAGACTGGTCTCTTTCTTGGATAAACCCACTCTATGCGCGGTCAITTTGGCGGTGCGCCCG 240
 QY 241 CAAGACTGCTAGCCAGTAGCTGCGGTTCGGAAGGCTTGTGTGTGTGTGTGTGTGTGTGT 300
 DB 241 CAAGACTGCTAGCCAGTAGCTGCGGTTCGGAAGGCTTGTGTGTGTGTGTGTGTGTGT 300
 QY 301 TGCCTTGGAGTGCCTCCCGGAGGTCTCTGTAGACCGTGCACATGAGCACAATCTTAAACC 360
 DB 301 TGCCTTGGAGTGCCTCCCGGAGGTCTCTGTAGACCGTGCACATGAGCACAATCTTAAACC 360
 QY 361 TCAAGAAAAAACCAAGAAAAACCAACCGTCCGCCCAACAAGACGTTAAGTTTCGCGCGG 420
 DB 361 TCAAGAAAAAACCAAGAAAAACCAACCGTCCGCCCAACAAGACGTTAAGTTTCGCGCGG 420
 QY 421 CGGCGAGATCGTTGGCGGAGTATCTTGTTCGCGCGAGGCGCCCGAGTTGGGTGTGG 480
 DB 421 CGGCGAGATCGTTGGCGGAGTATCTTGTTCGCGCGAGGCGCCCGAGTTGGGTGTGG 480
 QY 481 CGCGCAAGGAAGACTCTCGAGCGGTCCAGCAGCTGGAAGCGCGCAGCCCATCCCTAA 540
 DB 481 CGCGCAAGGAAGACTCTCGAGCGGTCCAGCAGCTGGAAGCGCGCAGCCCATCCCTAA 540
 QY 541 AGATCGGCGTCTCACTGGCAATCTGGGAAAAACAGGATACCCCTGCGCCCTATACGG 600
 DB 541 AGATCGGCGTCTCACTGGCAATCTGGGAAAAACAGGATACCCCTGCGCCCTATACGG 600
 QY 601 GAATCAGGAGCTCGGCTGGCAGGATGCTCTGTCCTCCCGCGAGGTTCCCGTCCCTCTTG 660
 DB 601 GAATCAGGAGCTCGGCTGGCAGGATGCTCTGTCCTCCCGCGAGGTTCCCGTCCCTCTTG 660
 QY 661 GGGCCCCAATGACCCCCCGGCATAGGTGCGCAACGTTGGTAAAGTCAATCGATACCCCTAAC 720

4081 GTATGCGCTCAGGGCTCAAAAGTCTAGTGTCTTAATCTCCTCGTGGCTGCCACCTGG 4140
 4141 GTTTGGGGCGTACTTGTCTAAGGACATGGGCATCAATCCCACTATTAGGACTGGAGTCA 4200
 4141 GTTTGGGGCGTACTTGTCTAAGGACATGGGCATCAATCCCACTATTAGGACTGGAGTCA 4200
 4201 GACTGTGACGACCGGGCGCCCATCAGTACTCCACATATGGCAATTCCTCGCGATGG 4260
 4201 GACTGTGACGACCGGGCGCCCATCAGTACTCCACATATGGCAATTCCTCGCGATGG 4260
 4261 GGGCTGCGGAGCGCGCCTATGACATCATATGCGATGAATGCGATGCGCGTGGACTC 4320
 4321 TACCACCATCTTGGCATCGGAACAGTCTTGTATCAAGCAGACAGACTGGGGTCAAGT 4380
 4321 TACCACCATCTTGGCATCGGAACAGTCTTGTATCAAGCAGACAGACTGGGGTCAAGT 4380
 4381 AACTGTGCTGGCTACAGCTACGCGCCCTGGGTCAAGTCAAGCAGACAGACTGGGGTCAAGT 4440
 4381 AACTGTGCTGGCTACAGCTACGCGCCCTGGGTCAAGTCAAGCAGACAGACTGGGGTCAAGT 4440
 4441 GAGGTGGCCCTTGGCAGAGGCGGAGATCCCTTCTATGGAGGCGGATTCCTCCCTGTC 4500
 4441 GAGGTGGCCCTTGGCAGAGGCGGAGATCCCTTCTATGGAGGCGGATTCCTCCCTGTC 4500
 4501 TTACATCAAGGAGGAGACATCTGATCTTCTGCGCATCAAGAAAAGTGTGACAGCT 4560
 4501 ATACATCAAGGAGGAGACATCTGATCTTCTGCGCATCAAGAAAAGTGTGACAGCT 4560
 4561 CGGGGGCCCTTGGGATGAGGCTTGAACAGTGGGATCAACAGAGGCTGGACGT 4620
 4561 CGGGGGCCCTTGGGATGAGGCTTGAACAGTGGGATCAACAGAGGCTGGACGT 4620
 4621 CTCCTGTAATACCAACTCAGGAGACGTAGTGGTCTGCGCCACCGACGCCCTCATGACAGG 4680
 4621 CTCCTGTAATACCAACTCAGGAGACGTAGTGGTCTGCGCCACCGACGCCCTCATGACAGG 4680
 4681 GTATCTGGGACCTTTGATCTCGTATGAGTCAAGCAGTGGGATCAAGTGGTGA 4740
 4681 GTTATTCTGGAGACTTTGATCTCGTATGAGTCAAGCAGTGGGATCAAGTGGTGA 4740
 4741 CTTTCACTTACCCACATTCACCATACCAACAGTGGTCTGCGCCACCGACGCCCTCATGACAGG 4800
 4741 CTTTCACTTACCCACATTCACCATACCAACAGTGGTCTGCGCCACCGACGCCCTCATGACAGG 4800
 4801 ACGTAGCCAGCGCGGGGTGCGACGGGTAGGGGAAGACTGGGCAATTTATAGGTATGTTTC 4860
 4801 ACGTAGCCAGCGCGGGGTGCGACGGGTAGGGGAAGACTGGGCAATTTATAGGTATGTTTC 4860
 4861 CACTGTGAGCGAGCCTCAGGAATGTTTGAAGTGTAGTGTCTCTGTGAGTGTCTACGACGC 4920
 4861 CACTGTGAGCGAGCCTCAGGAATGTTTGAAGTGTAGTGTCTCTGTGAGTGTCTACGACGC 4920
 4921 AGGGGGCGGATGTATGAGTCAACATCGGAGACCCCGTCAAGGCTCAGGCGTATTTT 4980
 4921 AGGGGGCGGATGTATGAGTCAACATCGGAGACCCCGTCAAGGCTCAGGCGTATTTT 4980
 4981 CAACACCGCGGTTTGGTGTGCGCAAGACCATCTTGAATTTTGGGAGGCGAGTTTTCAC 5040
 4981 CAACACCGCGGTTTGGTGTGCGCAAGACCATCTTGAATTTTGGGAGGCGAGTTTTCAC 5040
 5041 CGGCGCTCACACATAGTCCCACTTCTTCCCAAAACAGCAATCGGGGGAATTTT 5100
 5041 CGGCGCTCACACATAGTCCCACTTCTTCCCAAAACAGCAATCGGGGGAATTTT 5100
 5101 CGCATCTTAAAGCTTACAGGCTACAGTGTGGCTAGGGCCAAAGCCCGCCCGTC 5160
 5101 CGCATCTTAAAGCTTACAGGCTACAGTGTGGCTAGGGCCAAAGCCCGCCCGTC 5160
 5161 CTGGGAGCTCATGTGAAGTGTGATCTGAGCTCAAGCCCAACATCTGTTGGGCGCCACACC 5220

5161 CTGGGAGCTCATGTGAAAGTGTGATCTGACTCAAGCCCAACATCTGTTGGGCGCCACACC 5220
 5221 TCTCTGTACCGCTTGGGCTCTGTTTACCAACGAGGTCAACCTTCAACATCTCCGTGACGAA 5280
 5221 TCTCTGTACCGCTTGGGCTCTGTTTACCAACGAGGTCAACCTTCAACATCTCCGTGACGAA 5280
 5281 ATACATCTCCACCTCATGCTCAAGCGACTTGTAGGTCTATGACCAACATCTGTTGGC 5340
 5281 ATACATCTCCACCTCATGCTCAAGCGACTTGTAGGTCTATGACCAACATCTGTTGGC 5340
 5341 AGGGGAGTCTTGGGCGCGCTGCGCGCTATTTGCCCTGCGGACCGGGTGTGTTGATCAT 5400
 5341 TGGGGGGGTCTTGGGCGCGCTGCGCGCTATTTGCCCTGCGGACCGGGTGTGTTGATCAT 5400
 5401 CGGCGCGCTTGCACATTTAAACAGCAGCGCTGCTGCGCGCGCAACAGGAGTCTCTATGA 5460
 5401 CGGCGCGCTTGCACATTTAAACAGCAGCGCTGCTGCGCGCGCAACAGGAGTCTCTATGA 5460
 5461 GGGCTTTTGTATGAGATGGAGGAATGTGCTCTAGGGCGGCTCTCATTTGAAGAGGGGCGAG 5520
 5461 GGGCTTTTGTATGAGATGGAGGAATGTGCTCTAGGGCGGCTCTCATTTGAAGAGGGGCGAG 5520
 5521 GATAGCCGAGATGCTGAAGTCAAGATCAAGGCTTATTTGACGAAGCTTCCAAACAGC 5580
 5521 GATAGCCGAGATGCTGAAGTCAAGATCAAGGCTTATTTGACGAAGCTTCCAAACAGC 5580
 5581 TCAAGACATACAAACCTCTGTCAGGCTTTCATGGCCCAAGGTAGAGCAATTTCTGGGCCAA 5640
 5581 TCAAGACATACAAACCTCTGTCAGGCTTTCATGGCCCAAGGTAGAGCAATTTCTGGGCCAA 5640
 5641 ACACATGTGGAACTTTCATTTAGCGGCATCCAAATACCTCGCAGGACTATCAACACTGCCAGG 5700
 5641 ACACATGTGGAACTTTCATTTAGCGGCATCCAAATACCTCGCAGGACTATCAACACTGCCAGG 5700
 5701 GAACCTTGCAGTGTCTCATGATGGGCTTTCAGTCCCGCTTCAACAGTCCGCTGTCAAC 5760
 5701 GAACCTTGCAGTGTCTCATGATGGGCTTTCAGTCCCGCTTCAACAGTCCGCTGTCAAC 5760
 5761 AGACACACTATCTTCTCAACATTTTGGGGGCTGCTAGCATCCCAATTTGACACACC 5820
 5761 TAGCACCACTATCTTCTCAACATTTTGGGGGCTGCTAGCATCCCAATTTGACACACC 5820
 5821 CGCGGGGCGCACTGCTTCTGTTGTAGTGGGCTTTCAGTGGGAGCTGCGGTAGGAGTATAGG 5880
 5821 CGCGGGGCGCACTGCTTCTGTTGTAGTGGGCTTTCAGTGGGAGCTGCGGTAGGAGTATAGG 5880
 5881 CTTAGTAAAGTGTCTAGTGGACATCTTGGCAGGCTTATGTTGGGCGCATTTTGGGGGCTCT 5940
 5881 CTTAGTAAAGTGTCTAGTGGACATCTTGGCAGGCTTATGTTGGGCGCATTTTGGGGGCTCT 5940
 5941 CGTGCATTTCAAGATCATGCTGCGGAGAGCCCTTCAATGGAGGATGCTCAACTTGTCT 6000
 5941 CGTGCATTTCAAGATCATGCTGCGGAGAGCCCTTCAATGGAGGATGCTCAACTTGTCT 6000
 6001 GCTCGAATTTCTGCTTCTCGGGTGGCTTGTAGTGGGAGTCTCTGCGGCGCATTTCTGCG 6060
 6001 GCTCGAATTTCTGCTTCTCGGGTGGCTTGTAGTGGGAGTCTCTGCGGCGCATTTCTGCG 6060
 6061 CCGACAGTGGGACCGGGGAGGCGGCTTCCAAATGAGTGAATAGGCTCATTTGCGCTTTGC 6120
 6061 CCGACAGTGGGACCGGGGAGGCGGCTTCCAAATGAGTGAATAGGCTCATTTGCGCTTTGC 6120
 6121 TTCCAGAGGAATCAGCTCGCGCCCAACCTTACCTGAGCGGAGTGGATGCGTCCGAGCG 6180
 6121 TTCCAGAGGAATCAGCTCGCGCCCAACCTTACCTGAGCGGAGTGGATGCGTCCGAGCG 6180
 6181 TGTGACCAACTTCTTGGCTTCTTACCATTAACAGCTTGTGAGAGTCTCCAACTG 6240
 6181 TGTGACCAACTTCTTGGCTTCTTACCATTAACAGCTTGTGAGAGTCTCCAACTG 6240
 6241 GATTAATGAGGACTCGCCCATCTCCATGCGGCGGCTGTTGGTCTCCCGATGTTGGGACTG 6300
 6241 GATTAATGAGGACTCGCCCATCTCCATGCGGCGGCTGTTGGTCTCCCGATGTTGGGACTG 6300

QY	1921	CACCTACGTCGGGGAGAAATGAGACAGATGCTTCCCTATTGAACAGCACTCGACACC	1980	QY	3001	TGTGGTGTGTGACATAAACCAAGTGGCTCTTGGCGGTGCTTGGCCCTGCTTACCTCCTAAA	3060
DB	1921	CACCTACGTCGGGGAGAAATGAGACAGATGCTTCCCTATTGAACAGCACTCGACACC	1980	DB	3001	TGTGGTGTGTGACATAAACCAAGTGGCTCTTGGCGGTGCTTGGCCCTGCTTATCTCCTAAA	3060
QY	1981	GCTGGGTGTCATGCTTGGGTGTCAGTGGTGAATCTTCTTGGGTACACCAAGACTTGGG	2040	QY	3061	AGGTGCTTTTGACGCGGTGCCCTACTTGTGTGAGGCTCAGGCTCTACTGAGATGTGCAC	3120
DB	1981	GCAGGGTCTATGCTTGGGTGTCAGTGGTGAATCTTCTTGGGTACACCAAGACTTGGG	2040	DB	3061	AGGTGCTTTTGACGCGGTGCCCTACTTGTGTGAGGCTCAGGCTCTACTGAGATGTGCAC	3120
QY	2041	CGCACACCTGCTGCTAGAGCTGACTTCAAGCCGACAGGCTGCTTGTGCCCCAC	2100	QY	3121	CATGCAAGGCATCTCGCGGGGGGAGGTACGTCCAGATGGCGCTACTAGCCCTTGGCAG	3180
DB	2041	CGCACACCTGCTGCTAGAGCTGACTTCAAGCCGACAGGCTGCTTGTGCCCCAC	2100	DB	3121	CATGCAAGGCATCTCGCGGGGGGAGGTACGTCCAGATGGCGCTACTAGCCCTTGGCAG	3180
QY	2101	GGACTGTTTGGAAAGCATCCTGATACCACTTACCTCAAAATGCGGCTCTGGGCCCTGGGT	2160	QY	3181	GTGGAATGGCACTTATGACCACTTATGACCACTTATGACCGGCTTGTGAGTGGCTGTAGTGG	3240
DB	2101	GGACTGTTTGGAAAGCATCCTGATACCACTTACCTCAAAATGCGGCTCTGGGCCCTGGGT	2160	DB	3181	GTGGAATGGCACTTATGACCACTTATGACCACTTATGACCGGCTTGTGAGTGGCTGTAGTGG	3240
QY	2161	CAGCCAAAGTGGCTGATGACCTACCCCTACAGGCTCTGGCATTTACCCCTGACAGTTAA	2220	QY	3241	CTGCGGAGACTTGGCGGTGCGCGTGGAGCTTATCATCTTCACTTCACTGCGATGGAGAAAAGT	3300
DB	2161	CAGCCAAAGTGGCTGATGACCTACCCCTACAGGCTCTGGCATTTACCCCTGACAGTTAA	2220	DB	3241	CTGCGGAGACTTGGCGGTGCGCGTGGAGCTTATCATCTTCACTTCACTGCGATGGAGAAAAGT	3300
QY	2221	CTATACCATCTTCAAAATAGGATGTATGTTGGGGGGGTGAGCACAGGCTCACGGCTGC	2280	QY	3301	CATTTGCTGGGAGCGGAGACAGCTGCTTGTGGGGACATTTTACACGGACTTCCCGCTGTC	3360
DB	2221	CTATACCATCTTCAAAATAGGATGTATGTTGGGGGGGTGAGCACAGGCTCACGGCTGC	2280	DB	3301	CATTTGCTGGGAGCGGAGACAGCTGCTTGTGGGGACATTTTACACGGACTTCCCGCTGTC	3360
QY	2281	ATGCAATTTCACTCGTGGGATCGTTGCAACTTGGAGGACAGACAGAACTCAAAGTGC	2340	QY	3361	CGCCCGACTTGTGCGGAGGTCTCTTGGCCAGCTGATGCTATATCTTCAAGGGGTG	3420
DB	2281	ATGCAATTTCACTCGTGGGATCGTTGCAACTTGGAGGACAGACAGAACTCAAAGTGC	2340	DB	3361	CGCCCGACTTGTGCGGAGGTCTCTTGGCCAGCTGATGCTATATCTTCAAGGGGTG	3420
QY	2341	TCCTTTGTGCACTCCACCGAAATGGGCCATTTTACCTTGTCTTACTTCGACCTGCC	2400	QY	3421	GAGTCTTCTGGCCCCCATCACTGCTTACGCCACAGACAGACAGTGGCTTTTGGGCACCAT	3480
DB	2341	TCCTTTGTGCACTCCACCGAAATGGGCCATTTTACCTTGTCTTACTTCGACCTGCC	2400	DB	3421	GAGTCTTCTGGCCCCCATCACTGCTTACGCCACAGACAGACAGTGGCTTTTGGGCACCAT	3480
QY	2401	CGCTTGTCACTGCTTCTTCCACCTCCACCAAAACATCGTGGAGCTCAATTCATGTA	2460	QY	3481	AGTGGTGACATGACGGGGCGGACAGACAGAAAGGCTGGGGAAATTCAGGTCTCTGTC	3540
DB	2401	CGCTTGTCACTGCTTCTTCCACCTCCACCAAAACATCGTGGAGCTCAATTCATGTA	2460	DB	3481	AGTGGTGACATGACGGGGCGGACAGACAGAAAGGCTGGGGAAATTCAGGTCTCTGTC	3540
QY	2461	TGGCTTATCACTGCTTCAAAATACATGCTCCGATGGAGTGGGTAAATCTCTTATT	2520	QY	3541	CACAGTCACTCAGTCTTCTCTCGGAAATCCATCTCTCGGGGTCTTATGGACTGTACCA	3600
DB	2461	TGGCTTATCACTGCTTCAAAATACATGCTCCGATGGAGTGGGTAAATCTCTTATT	2520	DB	3541	CACAGTCACTCAGTCTTCTCTCGGAAATCCATCTCTCGGGGTCTTATGGACTGTACCA	3600
QY	2521	CTGCTCTTACGCGACCGAGGTGTCGCTCTTATGATGCTCATCTTGTGGGCCA	2580	QY	3601	TGGAGCTGCGAACAGACTCTGCGGGCTCAGGGGTCCGCTCAGCAGATGTACTCCAG	3660
DB	2521	CTGCTCTTACGCGACCGAGGTGTCGCTCTTATGATGCTCATCTTGTGGGCCA	2580	DB	3601	TGGAGCTGCGAACAGACTCTGCGGGCTCAGGGGTCCGCTCAGCAGATGTACTCCAG	3660
QY	2581	GGCGGACGACACTAGAGAGCTGCTATCTTGCAGCTGCGAGCGCAGCTAGCTGCAA	2640	QY	3661	TGCTAGGGGGACTTATGAGGGTGGCCAGCCCCCTGGGACTTAAATCTTTGAGGCGGTG	3720
DB	2581	GGCGGACGACACTAGAGAGCTGCTATCTTGCAGCTGCGAGCGCAGCTAGCTGCAA	2640	DB	3661	TGCTAGGGGGACTTATGAGGGTGGCCAGCCCCCTGGGACTTAAATCTTTGAGGCGGTG	3720
QY	2641	TGGCTTCTATATTTTGTCTATCTTTTGTGGCTGCTTGTGATCATCAAGGGTGGGTAGT	2700	QY	3721	CACGTGTGAGGGTGCACCTGTACTGTGTACGCGGAAACGCTGTATCTCCCGGTCTG	3780
DB	2641	TGGCTTCTATATTTTGTCTATCTTTTGTGGCTGCTTGTGATCATCAAGGGTGGGTAGT	2700	DB	3721	CACGTGTGAGGGTGCACCTGTACTGTGTACGCGGAAACGCTGTATCTCCCGGTCTG	3780
QY	2701	CCCTTACGCTTATCTTCTGCTGCTGCTTGTGATGCTTGTGCTTCTCTAGCAAT	2760	QY	3781	AAGACGGGGGACAAACGGGGAGCGCTACTCTCCCGGAGACCTTTTCCACCTTGAAGGG	3840
DB	2701	CCCTTACGCTTATCTTCTGCTGCTGCTTGTGATGCTTGTGCTTCTCTAGCAAT	2760	DB	3781	AAGACGGGGGACAAACGGGGAGCGCTACTCTCCCGGAGACCTTTTCCACCTTGAAGGG	3840
QY	2761	GGCCCAACGCTTATGCTTATGACCATCTGTGCTGATGGCCAGATAGGAGGCTGCT	2820	QY	3841	GTCTCAGAGGGCCCGGTGCTATGCCCGAGGGGACAGCTGTGCGAGTCTTCCCGGGCAGC	3900
DB	2761	GGCCCAACGCTTATGCTTATGACCATCTGTGCTGATGGCCAGATAGGAGGCTGCT	2820	DB	3841	GTCTCAGAGGGCCCGGTGCTATGCCCGAGGGGACAGCTGTGCGAGTCTTCCCGGGCAGC	3900
QY	2821	GGTAATGATCACTCTTTTACTCTCACCCCGGGTATAAGACCTTCTCAGCGGTTT	2880	QY	3901	TGTGTCTCTCGGGGGCTGCTAAGTCCATAGATTTTATCTCCCGTTGAGACACTGACAT	3960
DB	2821	GGTAATGATCACTCTTTTACTCTCACCCCGGGTATAAGACCTTCTCAGCGGTTT	2880	DB	3901	TGTGTCTCTCGGGGGCTGCTAAGTCCATAGATTTTATCTCCCGTTGAGACACTGACAT	3960
QY	2881	GTGGTGTGTGCTATCTTCTGACCTTACCTGAGGAGCTATGGTCCAGGATGGGCACACC	2940	QY	3961	CGTACCGGCTCCCCACCTTTAGTGAACAACAGCACACACCTGCTGTGCCCCAACTTA	4020
DB	2881	GTGGTGTGTGCTATCTTCTGACCTTACCTGAGGAGCTATGGTCCAGGATGGGCACACC	2940	DB	3961	CGTACCGGCTCCCCACCTTTAGTGAACAACAGCACACACCTGCTGTGCCCCAACTTA	4020
QY	2941	TATGAGGTGGGGTGGCTGATGGCATCATATGGGGCGTGGCATATTTCTACCCAGG	3000	QY	4021	TCAGGTGGGTACTTTCATGCCCCGACTGTCAGTGGAAAGAGACCAAGTCTCTGTGCG	4080
DB	2941	TATGAGGTGGGGTGGCTGATGGCATCATATGGGGCGTGGCATATTTCTACCCAGG	3000	DB	4021	TCAGGTGGGTACTTTCATGCCCCGACTGTCAGTGGAAAGAGACCAAGTCTCTGTGCG	4080
				QY	4081	ATATGTGCTCAGGGGTATAAAGTGTAGTGTCTTAAATCCCTCAGTGGCTGCCACCCCTGGG	4140

Copied from 09540843 on 05/19/2004

```

8881 CAATTATGGCTAGGAAACATCATCCAGTACGCCCAACCATATATGGGCTCGCATGTCCT 8940
8941 GATGACACACTTCTTCTCCATCTCATGGCTCAAGACACGCTGGACACAGCACTCAACTT 9000
8941 GATGACACACTTCTTCTCCATCTCATGGCTCAAGACACGCTGGACACAGCACTCAACTT 9000
9001 TGAGATGTACGAGCGGTGTACTCCGTGAGTCCCTTGGACCTCCAGCCATAATTAAG 9060
9001 TGAGATGTACGAGCGGTGTACTCCGTGAGTCCCTTGGACCTCCAGCCATAATTAAG 9060
9061 GTTACATGGGCTTGAGCGCTTCTCTGCACATACATCTCCCAAGCACTGACCGGT 9120
9061 GTTACATGGGCTTGAGCGCTTCTCTGCACATACATCTCCCAAGCACTGACCGGT 9120
9121 GGCCTTCAGCCCTCAGAAAACTTGGGGCGCCACCCCTCAGAGGTGGAAGAGCGGCGACG 9180
9121 GGCCTTCAGCCCTCAGAAAACTTGGGGCGCCACCCCTCAGAGGTGGAAGAGCGGCGACG 9180
9181 TGAGTTCAGGGCGTCCCTCATCTCCGTGGGGGAGAGCGGCGTTCGGTGCATATCT 9240
9181 TGAGTTCAGGGCGTCCCTCATCTCCGTGGGGGAGAGCGGCGTTCGGTGCATATCT 9240
9241 CTTCAATTCGGCGGTGAAGCAAGCTCAAACTCACTCCATTCGCCGGAAGCGGCTCCT 9300
9241 CTTCAATTCGGCGGTGAAGCAAGCTCAAACTCACTCCATTCGCCGGAAGCGGCTCCT 9300
9301 GGATTTATCCAGCTGGTTCACCGTCCGGCGCGGGGCGGCGACATTTATCAGCGCTGTC 9360
9301 GGATTTATCCAGCTGGTTCACCTGTCGGCGCGGGGCGGCGACATTTATCAGCGCTGTC 9360
9361 GCGTCCCGACCCCGCTTATGCTCTTTGGCTTACCTCTCTTTTGTAGGGGTAGGCTT 9420
9361 GCGTCCCGACCCCGCTTATGCTCTTTGGCTTACCTCTCTTTTGTAGGGGTAGGCTT 9420
9421 TTTCTACTCCCGCTCGGTAGAGCGGACACATTTAGCTACACTCCATAGCTAACTGTC 9480
9421 TTTCTACTCCCGCTCGGTAGAGCGGACACATTTAGCTACACTCCATAGCTAACTGTC 9480
9481 CTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 9540
9481 CTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 9540
9541 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 9588
9541 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 9588

```

RESULT 7

```

E07362
LOCUS
DEFINITION
E07362
CDNA of Hepatitis non-A non-B virus.
E07362
ACCESSION
E07362.1
GI:2175501
E07362.1
VERSION
E07362.1
KEYWORDS
E07362.1
SOURCE
E07362.1
ORGANISM
E07362.1
REFERENCE
E07362.1
AUTHORS
E07362.1
TITLE
E07362.1
JOURNAL
E07362.1
COMMENT
E07362.1

```

QY	5641	ACACATGTGGAACTTCAATAGCGGCATCAATACCTCGCAGACTATCAACACTGCCAGG	5700	6721	CTGGGTGACCGAGTGCAGATCCATAGGTTTGGCCCCACACGGAAGCGTTTTCCTCGGA	6780
DB	5641	ACACATGTGGAACTTCAATAGCGGCATCAATACCTCGCAGACTATCAACACTGCCAGG	5700	6721	CTGGGTGACCGAGTGCAGATCCATAGGTTTGGCCCCACACGGAAGCGTTTTCCTCGGA	6780
QY	5701	GAACCTGCACTAGTTCATATGATGCGTTCAGTCCGCGCTCAACAGTCCGCTGCAAC	5760	6781	TCAGGTCTCGTTCGCGTTCGCGTTCGCGTTCGCGTTCGCGTTCGCGTTCGCGTTCGCG	6840
DB	5701	GAACCTGCACTAGTTCATATGATGCGTTCAGTCCGCGCTCAACAGTCCGCTGCAAC	5760	6781	TCAGGTCTCGTTCGCGTTCGCGTTCGCGTTCGCGTTCGCGTTCGCGTTCGCGTTCGCG	6840
QY	5761	AAGCACCATCTCTCTCAACATTTTGGGGGCTGGCTAGCATCCCAATTTGCAACACC	5820	6841	CCCTGAAACCCGACACAGACGATTTGATCTCATGTCTAAACAGATCCATCTCATATCACGC	6900
DB	5761	AAGCACCATCTCTCTCAACATTTTGGGGGCTGGCTAGCATCCCAATTTGCAACACC	5820	6841	CCCTGAAACCCGACACAGACGATTTGATCTCATGTCTAAACAGATCCATCTCATATCACGC	6900
QY	5821	CGCGGGGCGACCTGCTTCGTTGTCAGTGGCTTCAGTGGGAGCTCCGCTAGGAGTATAGG	5880	6901	GGAGACTGCAGCGCGCGCTTTAGCGCGGGGTTCACCCCATCCGAGGCAAGCTCTCGGC	6960
DB	5821	CGCGGGGCGACCTGCTTCGTTGTCAGTGGCTTCAGTGGGAGCTCCGCTAGGAGTATAGG	5880	6901	GGAGACTGCAGCGCGCGCTTTAGCGCGGGGTTCACCCCATCCGAGGCAAGCTCTCGGC	6960
QY	5881	CTTAGGTAAAGTGTCTAGTGGACATCTCGGAGGGTATGGTGGGGGCTCTTCGCGGCTCT	5940	6961	GAGCAGCTATCGGCACCATCGCTCGAGCCACCTGACCCACCCACCGGCAAGCTCTCAGC	7020
DB	5881	CTTAGGTAAAGTGTCTAGTGGACATCTCGGAGGGTATGGTGGGGGCTCTTCGCGGCTCT	5940	6961	GAGCAGCTATCGGCACCATCGCTCGAGCCACCTGACCCACCCACCGGCAAGCTCTCAGC	7020
QY	5941	CGTGGCATTCAGATCATGTCTGGCGAGAGCCCTCCATGGAGGATGTTGTCAACCTGCT	6000	7021	TGTGGACATGTTGATGCTTAACCTCTTCATCGGGGGCGATGTGACTCGGATAGTCTGG	7080
DB	5941	CGTGGCATTCAGATCATGTCTGGCGAGAGCCCTCCATGGAGGATGTTGTCAACCTGCT	6000	7021	TGTGGACATGTTGATGCTTAACCTCTTCATCGGGGGCGATGTGACTCGGATAGTCTGG	7080
QY	6001	GCCTGGAAATTCCTCTCCGGGCTCTTGTGTAGTGGGAGTCACTCGCGCGCCATTCTGG	6060	7081	GTCCAAAGTGTCTTCGAGCTCTCTCGACCCCAATGTTGGAATCGGAAAGAGGACCTTGA	7140
DB	6001	GCCTGGAAATTCCTCTCCGGGCTCTTGTGTAGTGGGAGTCACTCGCGCGCCATTCTGG	6060	7081	GTCCAAAGTGTCTTCGAGCTCTCTCGACCCCAATGTTGGAATCGGAAAGAGGACCTTGA	7140
QY	6061	CGGACAGTGGGACCGGGGAGGGCGCTGTCCAATGGATGAATAGACTCATTCGCTTTGC	6120	7141	GCCTTCGATACATCATGAAATACATGCTCCCAAGAGAGGTTTCCACAGCTTTACCGGC	7200
DB	6061	CGGACAGTGGGACCGGGGAGGGCGCTGTCCAATGGATGAATAGACTCATTCGCTTTGC	6120	7141	GCCTTCGATACATCATGAAATACATGCTCCCAAGAGAGGTTTCCACAGCTTTACCGGC	7200
QY	6121	TTCCAGAGGAATACAGTCCGCGCCCAACCTACGTACGAGTGGAGTGGAGTGGAGTGG	6180	7201	CTGGGACGCGCTGATTACAAACCCACCGCTTCTGGAATCGTGGAAAGAGGACGATTACA	7260
DB	6121	TTCCAGAGGAATACAGTCCGCGCCCAACCTACGTACGAGTGGAGTGGAGTGGAGTGG	6180	7201	CTGGGACGCGCTGATTACAAACCCACCGCTTCTGGAATCGTGGAAAGAGGACGATTACA	7260
QY	6181	TGTGACCCAACTACTTGGCTCCCTTACCAATACAGGCTGCTCAGAGACTCCACAACTG	6240	7261	ACCGGCGCTGTTGGCGGCTGCTCTCCCTCTAGGAAACCCCGAGCCCTCCGCC	7320
DB	6181	TGTGACCCAACTACTTGGCTCCCTTACCAATACAGGCTGCTCAGAGACTCCACAACTG	6240	7261	ACCGGCGCTGTTGGCGGCTGCTCTCCCTCTAGGAAACCCCGAGCCCTCCGCC	7320
QY	6241	GATTTACTGAGGACTGCCCATCCATCGCGCGGCTCGTGGCTCCGCGATGTGTGGGACTG	6300	7321	AAGGAGCGCGGACAGTGGGCTTAAGTGAAGACTTCATAGAGAGTGGCTTTCAACAGCT	7380
DB	6241	GATTTACTGAGGACTGCCCATCCATCGCGCGGCTCGTGGCTCCGCGATGTGTGGGACTG	6300	7321	AAGGAGCGCGGACAGTGGGCTTAAGTGAAGACTTCATAGAGAGTGGCTTTCAACAGCT	7380
QY	6301	GGTTTGACCATCTTAAACAGACTTTAAACCTGGCTGACCTCCAAATTTGTCCCAAGAT	6360	7381	GGCCATTAAATGCTTTGGCGCGCCCGCCCGGCGGATTCAGGCTTTTCCACGGGGGC	7440
DB	6301	GGTTTGACCATCTTAAACAGACTTTAAACCTGGCTGACCTCCAAATTTGTCCCAAGAT	6360	7381	GGCCATTAAATGCTTTGGCGCGCCCGCCCGGCGGATTCAGGCTTTTCCACGGGGGC	7440
QY	6361	GCCCGGCTCCCTTTGCTCTCTGTCGAAAGGGGTACAGGGGCTGGTGGCGGCACTGG	6420	7441	GGGCGCTCCGATTCGCGGAGTCAAGCGCTCTCTGATGATTTGGGCTTTTCGAGACAGG	7500
DB	6361	GCCCGGCTCCCTTTGCTCTCTGTCGAAAGGGGTACAGGGGCTGGTGGCGGCACTGG	6420	7441	GGGCGCTCCGATTCGCGGAGTCAAGCGCTCTCTGATGATTTGGGCTTTTCGAGACAGG	7500
QY	6421	CATCATGACACACGCTGCTTCGCGCGGCAATATCTCTGGCAATGCTCGGCTTGGGCTC	6480	7501	TTCCATCTCTTCATGCCCCCTCTCGAGGGGAGTTCGAGATCCAGACTTCGAGCCTGA	7560
DB	6421	CATCATGACACACGCTGCTTCGCGCGGCAATATCTCTGGCAATGCTCGGCTTGGGCTC	6480	7501	TTCCATCTCTTCATGCCCCCTCTCGAGGGGAGTTCGAGATCCAGACTTCGAGCCTGA	7560
QY	6481	CATGAGAAATCAACGGGCTTAAGACCTGATGAATATCTGGAGGAGGACCTTTCTATCAA	6540	7561	GCAGGTAGAGCCCAACCCCGGCGGCGGCTCGAGCTCCCGGCTCGGCTCGGCTCGG	7620
DB	6481	CATGAGAAATCAACGGGCTTAAGACCTGATGAATATCTGGAGGAGGACCTTTCTATCAA	6540	7561	GCAGGTAGAGCCCAACCCCGGCGGCGGCTCGAGCTCCCGGCTCGGCTCGGCTCGG	7620
QY	6541	TTGTTACAGGGGCGAGTGGTGGCGGCAACCCCGGCGGCAACTTTAAGTGGCCTATCG	6600	7621	GTCTGTGCTACTTCTCGAGGAGGAGCTCCGCTCGTGTGCTCTCTCATGCTACTCTC	7680
DB	6541	TTGTTACAGGGGCGAGTGGTGGCGGCAACCCCGGCGGCAACTTTAAGTGGCCTATCG	6600	7621	GTCTGTGCTACTTCTCGAGGAGGAGCTCCGCTCGTGTGCTCTCTCATGCTACTCTC	7680
QY	6601	GAGGCTGGGCGCTCAGAGTACCGGAGGTGACGAGCAGCGGCTCATACCTATACAAAC	6660	7681	CTGGAACCGGGCTTAATACTCTCTGTTAGTCCGAGAGGAGGATTTACCGATTAAACC	7740
DB	6601	GAGGCTGGGCGCTCAGAGTACCGGAGGTGACGAGCAGCGGCTCATACCTATACAAAC	6660	7681	CTGGAACCGGGCTTAATACTCTCTGTTAGTCCGAGAGGAGGATTTACCGATTAAACC	7740
QY	6661	AGGACTTACCATGATACTTGAAGTTCCTTCCCACTACCTTCTCCAGAGTCTTTTTC	6720	7741	CTTGAGCAACTCCCTGTTGGCGATATCAACAAGGTGTACTGTACCAACAAGAGGCG	7800
DB	6661	AGGACTTACCATGATACTTGAAGTTCCTTCCCACTACCTTCTCCAGAGTCTTTTTC	6720	7741	CTTGAGCAACTCCCTGTTGGCGATATCAACAAGGTGTACTGTACCAACAAGAGGCG	7800
				7801	CTCACTAAGGGCTAAAGAGTAACTTTTGTAGGATGCAAGTGTCTGCTCTCTACTACGA	7860

Db 3421 |||||GAGTCTTTCGCGCCCATCATCTGTTATGCCAGCAGACACGCGCGCTTTTGGGCGACCAT 3480
 Qy 3481 AGTGGTGAGCATGACCGGGCGGCACAGACAGACAGGCTGGGGAATTCAGGTCTGTGC 3540
 Db 3481 AGTGGTGAGCATGACCGGGCGGCACAGACAGACAGGCTGGGGAATTCAGGTCTGTGC 3540
 Qy 3541 CACAGTCACTCAGTCTCTTCGGAACATCCATCTCGGGGTTTTTGTGGACTGTCTACCA 3600
 Db 3541 CACGGTCACTCAGTCTCTTCGGAACAACTCTCGGGGTTCTTATGGACTGTCTACCA 3600
 Qy 3601 TGAGCTGGCAACAAGACTCTGGCGCGCTCACGGGGTTCGGGTACGAGATGTACTCCAG 3660
 Db 3601 TGAGCTGGCAACAAGACTCTAGCCGGCTCACGGGGTTCGGGTACGAGATGTACTCCAG 3660
 Qy 3661 TGCTGAGGGGACTTGTAGTGGGTGGCCAGCCCGCTGGGACTAAATCTTTGAGCGGTG 3720
 Db 3661 TGCTGAGGGGACTTGTAGTGGGTGGCCAGCCCGCTGGGACTAAATCTTTGAGCGGTG 3720
 Qy 3721 CAGCTGTGGAGCGGTTCGACCTGTACCTGTGTCAACGGGAACGCTGTCTATCCCGGTGC 3780
 Db 3721 CAGCTGTGGAGCGGTTCGACCTGTATCTGTGTCAACGGGAACGCTGTCTATCCCGGTGC 3780
 Qy 3781 AAGACGGGGGCAAAACGGGAGCGCTACTCTCTCCCGAGACCTCTTTCCACCTTGAAGG 3840
 Db 3781 AAGACGGGGGCAAAACGGGAGCGCTACTCTCTCCCGAGACCTCTTTCCACCTTGAAGG 3840
 Qy 3841 GTCTCAGGAGGCGCGTGTATGCCCGGCGCCAGCTCGGAGTCTTCGGGCGAGC 3900
 Db 3841 GTCTCAGGAGGCGCGTGTATGCCCGGCGCCAGCTCGGAGTCTTCGGGCGAGC 3900
 Qy 3901 TGTGTGCTCTCGGGGCTGGCTAAGTCCATAGATTTTCATCCCGGTTGAGACACTCCGACAT 3960
 Db 3901 CGTGTGCTCTCGGGGCTGGCGAAGTCCATAGATTTTCATCCCGGTTGAGACACTTGACAT 3960
 Qy 3961 CGTCAACGGGTCTCCCACTTTTGTAGTGAACAAGCAGACCACTCTGTGCGCCCAACTTA 4020
 Db 3961 CGTCACTCGGTCTCCCACTTTTGTAGTGAACAAGCAGACCACTCTGTGCGCCCAACTTA 4020
 Qy 4021 TCAGTGGGTACTTTCATGCCCGGCTGGTAGTGGGAAGACCAAAAGTCTCTGTGCGC 4080
 Db 4021 TCAGTGGGTACTTTCATGCCCGGCTGGTAGTGGGAAGACCAAAAGTCTCTGTGCGC 4080
 Qy 4081 ATATGCTGCTCAGGGGTATAAGTGTAGTGTCTTAATCCCTCAGTGGCTGCCACCTGGG 4140
 Db 4081 GTATGCCGCTCAGGGGTACAAAGTGTAGTGTCTTAATCCCTCAGTGGCTGCCACCTGGG 4140
 Qy 4141 GTTTGGGGGTACTTGTCTAAGGCACATGGCATCAATCCCAACATTTAGGACTGGGGTCAG 4200
 Db 4141 GTTTGGGGGTACTTGTCTAAGGCACATGGCATCAATCCCAACATTTAGGACTGGGGTCAG 4200
 Qy 4201 GACTGTGACGACCGGGGCGCCATCAGTACTCCACATATGGCAAAATTCCTCGCGATGG 4260
 Db 4201 GACTGTGACGACCGGGGCGCCATCAGTACTCCACATATGGCAAAATTCCTCGCGATGG 4260
 Qy 4261 GGGCTGTGCGGGGCGGCTCAGACATCATATGTGATGAATGCCATGCCGTGGACTC 4320
 Db 4261 GGGCTGTGCGGGGCGGCTCAGACATCATATGTGATGAATGCCATGCCGTGGACTC 4320
 Qy 4321 TACACCATCTCTTGGCATTCGGACAGCTCTTGTATCAAGCAGAGACAGCTGGGGTCAGACT 4380
 Db 4321 TACACCATCTCTCGCATTCGGACAGCTCTTGTATCAAGCAGAGACAGCTGGGGTCAGACT 4380
 Qy 4381 AACTGTCTGGCTACAGTACGCGCTCGGTCAGTCAACACCCCGCCACCCACACATAGA 4440
 Db 4381 AACTGTCTGGCTACAGTACGCGCTCGGTCAGTCAACACCCCGCCACCCACACATAGA 4440
 Qy 4441 GGAGTGGCCCTTGGGAGAGGCGGAGATCCCTTCTATGGAGGCGGATTCCTCTGTGC 4500
 Db 4441 GGAGTGGCCCTTGGGAGAGGCGGAGATCCCTTCTATGGAGGCGGATTCCTCTGTGC 4500
 Qy 4501 TTATCATCAAGGAGGAGACATCTGATCTTCTGTCCATTTCAAGAAAGTGTGACGAGCT 4560
 Db 4501 TTATCATCAAGGAGGAGACATCTGATCTTCTGTCCATTTCAAGAAAGTGTGACGAGCT 4560

Db 4501 ATACATCAAGGAGGAGAGACATTTGATCTTCTGCCACTCAAGAAAAAGTGTGACGAGCT 4560
 Qy 4561 CGCGGCGGCTTTCGGGGTATGGGGTGTGAACCTCAGTGGCATACTACAGAGGGTGTGACGCT 4620
 Db 4561 CGCGGCGGCTTTCGGGGTATGGGGTGTGAACGCACTGGCATACTACAGAGGGTGTGACGCT 4620
 Qy 4621 CTCCTGTAATCAAACTCAGGGAGACGTAGTGGTGTGCGCACCGAGCGCTCTATGACAGG 4680
 Db 4621 CTCCTGTAATCAAACTCAGGGAGACGTAGTGGTGTGCGCACCGAGCGCTCTATGACGGG 4680
 Qy 4681 GTATACCTGGGACTTTTGAATCGGTGATCGACTGCAACGTAGCGGTCACTCAAGTGTGAGA 4740
 Db 4681 GTTACTTGAGACTTTTGAATCGGTGATCGACTGCAACGTAGCGGTCACTCAAGTGTGAGA 4740
 Qy 4741 CTTTCACTTTAGACCCCACTTCAACATAACCAACACAGACTGTCCCTCAAGACGCTGTCTC 4800
 Db 4741 CTTTCACTTTGACCCCACTTCAACATAACCAACACAGACTGTCCCTCAAGACGCTGTCTC 4800
 Qy 4801 AGCTAGCCAGCGCGGGTTCGCAACGGGTAGGGAGAGACTGGGCAATTTATAGGTATGTTTC 4860
 Db 4801 AGCTAGCCAGCGCGGGTTCGCAACGGGTAGGGAGAGACTGGGTAATTTATAGGTATGTTTC 4860
 Qy 4861 CACTGTGTAGCGGCTCAGGAAATGTTTGCAGTGTAGTGTCTGTGAGTGTCTTACGACGC 4920
 Db 4861 CACTGTGTAGCGGCTCAGGAAATGTTTGCAGTGTAGTGTCTGTGAGTGTCTTACGATGC 4920
 Qy 4921 AGGGGCGCATGTGTAGTGTACACATCGAGAGACACCGTCAGGCTCAGGGCGGTATTT 4980
 Db 4921 AGGGGCGCATGTGTATGAGTGTACACAGGGAGACACCGTCAGGCTCAGAGCATATTT 4980
 Qy 4981 CAAACGCGCGTTCCTGTGTGCGCAAGACCATCTTGAATTTTGGGAGGAGTGTTCAC 5040
 Db 4981 CAAACACCTCTGTTTCCCTGTGTGCGCAAGACCATCTTGAATTTTGGGAGGAGTGTTCAC 5040
 Qy 5041 CGGCTCTCACACATAGATGCCACTTCTTCCCAAAACAAAGCAATTCGGGGGAAATTT 5100
 Db 5041 CGGCTCTCACACATAGATGCCACTTCTTCCCAAAACAAAGCAATTCGGGGGAAATTT 5100
 Qy 5101 CGCATCTTAAACAGCCTTACAGGCTACAGTGTGCGTGTAGGGCCAAAGCCCCCGGTC 5160
 Db 5101 CGCATCTTAAACAGCCTTACAGGCTACAGTGTGCGTGTAGGGCCAAAGCCCCCGGTC 5160
 Qy 5161 CTGGGAGCTCATGTGGAAGTGTGACTGCACTCAAGCCCACTCTGTGGGCCCCACACC 5220
 Db 5161 CTGGGAGCTCATGTGGAAGTGTGACTGCACTCAAGCCCACTCTGTGGGCCCCACACC 5220
 Qy 5221 TCTCTGTACCGCTTGGGCTCTGTACCAAGGCTCACCTCACACATCCCTGTGAGCA 5280
 Db 5221 TCTCTGTACCGCTTGGGCTCTGTACCAAGGCTCACCTCACACATCCCTGTGAGCA 5280
 Qy 5281 ATACATCGCCACCTGCATGCAAGCCGACCTTGAAGTGTATGACCAAGCACTGGGCTTTGGC 5340
 Db 5281 ATACATCGCCACCTGCATGCAAGCCGACCTTGAAGTGTATGACCAAGCACTGGGCTTTAGC 5340
 Qy 5341 AGGGGAGTCTTGGCGCGCTCGCCCGTATTCCTGTGGGAGCCGGGTGTGTTGCAATCAT 5400
 Db 5341 TGGGGGGGCTTGTGGCGCGCTCGCCCGTACTTCTGTGGGAGCCGGGTGTGTTGCAATCAT 5400
 Qy 5401 CGGCGCTTGCACATTAACAGCGCGCTGTGCGCGGAGCAAGGAGTCTCTTATGA 5460
 Db 5401 CGGCGCTTGCACATTAACAGCGCGCTGTGCGCGGAGCAAGGAGTCTCTTATGA 5460
 Qy 5461 GGCTTTTGTAGTGAATGGAGAAATGTGCTCTAGGGGGGTCTCTCAATGAAGAGGGGAGCG 5520
 Db 5461 GGCTTTTGTAGTGAATGGAGAAATGTGCTCTAGAGCGGCTCTCAATGAAGAGGGGAGCG 5520
 Qy 5521 GATAGCGAGATCTGAAAGTCCAAGATCCAAGGCTTTATTTGAGCAAGCTTCCAAACAGC 5580
 Db 5521 GATAGCGAGATCTGAAAGTCCAAGATCCAAGGCTTTATTTGAGCAAGCTTCCAAACAGC 5580
 Qy 5581 TCAAGACATACACCACTGTGAGGCTTCACTGGCCCAAGGTAGAACAAATTTCTGGGCGAA 5640
 Db 5581 TCAAGACATACACCACTGTGAGGCTTCTTGGCCCAAGGTAGAACAAATTTCTGGGCGAA 5640

2341	TCCTTTGTTGACATCCACCAACGGAATGGGCCATTTTACCTTGCTCTTACTCGACCTGCC	2400
2341		2400
2341	TCCTTTGCTGACATCCACACGAGATGGGCCATTTTACCTTGCTCTTACTCGACCTGCC	2400
2341		2400
2401	CGCTTGTGCGACTGGTCTTCTCCACCTCCACCAAAACATCGTGGACGTACAAATCATGTA	2460
2401		2460
2401	CGCTTGTGCGACTGGTCTTCTCCACCTCCACCAAAACATCGTGGACGTACAAATCATGTA	2460
2401		2460
2461	TGGCTATCACTCGCCCTCACAAATACATCGTCCGATGGGAGTGGGTAATCTCTTAAT	2520
2461		2520
2521	CCTGCTTTAGCGACGCGCAGGGTTTGGCCTGCTTATGAAATGCTCATCTTTGTTGGGCCA	2580
2521		2580
2521	CCTGCTTTAGCGACGCGCAGGGTTTGGCCTGCTTATGAAATGCTCATCTTTGTTGGGCCA	2580
2521		2580
2581	GGCCGAGCAGCACTAGAGAGCTGGTCACTCTTGACGCTGGAGCGCAGCTAGCTGCAA	2640
2581		2640
2581	GGCCGAGCAGCACTAGAGAGTGGTGGTCTTGACGCTGGAGCGCAGCTAGCTGCAA	2640
2581		2640
2641	TGGCTTCCTATATTTTGTGCATCTTTTTCGTGGCTGCTTGGTACATCAAGGCTCGGTAAGT	2700
2641		2700
2641	TGGCTTCCTATATTTTGTGCATCTTTTTCGTGGCTGCTTGGTACATCAAGGCTCGGTAAGT	2700
2641		2700
2701	CCCTTTAGCTACCTATTCCTCTCACTGGCCTGTGCTTCCTTTAGCTCTCTCTAGCAAT	2760
2701		2760
2701	CCCTTTAGCTACCTATTCCTCTCACTGGCCTATGGTCCCTTTGGCCTACTGCTCCTAGCAAT	2760
2701		2760
2761	GGCCGACAGGCTTATGCTTATGACGATCTGTGCATGCGCAGATAGGAGCGGCTCTGCT	2820
2761		2820
2761	GGCCGACAGGCTTATGCTTATGACGATCTGTGCATGCGTCCAGATAGGAGCAGCTCTGTT	2820
2761		2820
2821	GGTAATGATCACTCTCTTTACTCTCACCCCGGGTATAAGACCCTTCTCACCCGGTTTTT	2880
2821		2880
2821	GGTACTGATCACTCTCTTTTACACTCACCCCGGGTATAAGACCCTTCTCACCCGGTTTTCT	2880
2821		2880
2881	GTGTGGTTGTGCTATCTTCTGACCTCTGGGGAAAGCTATGTGTCCAGGAGTGGGCACCA	2940
2881		2940
2881	GTGTGGTTGTGCTATCTTCTGACCTCTGGCGAAGCTATGGTCCAGGAGTGGGCACCA	2940
2881		2940
2941	TATCAGAGTGGCGGTGGCGGTGATGGGATCATATGGGCGGTGGCGATATTCTACCGAG	3000
2941		3000
2941	TATCAGAGTGGCGGTGGCGGTGATGGGATCATATGGGCGGTGGCGATATTCTGCCCGGG	3000
2941		3000
3001	TGTGTGTTTACATAAACAAATGCGCTCTTGGCGGTGCTTGGCGCTGCTTACCTCTTAAA	3060
3001		3060
3001	TGTGTGTTTACATAAACAAATGCGCTCTTGGCGGTGCTTGGCGCTGCTTATCTCTTAAA	3060
3001		3060
3061	AGGTGCTTTGACGCGGTGGCGTACTTCTGTCCAGGGCTCACGCTCTACTGAGGATGTGCAC	3120
3061		3120
3061	AGGTGCTTTGACGCGGTGGCGTACTTCTGTCCAGGGCTCACGCTCTACTAAGGATGTGCAC	3120
3061		3120
3121	CATGGCAAGGATCTCGGGGGGCGAGTACGTTCCAGATGGCGCTACTAGGCCCTTGGCAG	3180
3121		3180
3121	CATGGTAAGGATCTCGCGGGGGTAGGTACGTCCAGATGGTGTCTACTAGCCCTTGGCAG	3180
3121		3180
3181	GTGACTTGGCACTTTACATCTATGACCACTCACCCCTATGTGCGAATTTGGGCTGTAGTGG	3240
3181		3240
3181	GTGACTTGGCACTTTACATCTATGACCACTCACCCCTATGTGCGAATTTGGGCTGTAAATGG	3240
3181		3240
3241	CCTCGGGACCTGCGGGTTCGCGTTGAGCCTTATCATCTTTCAGTCCGATGGAGAGAAAGT	3300
3241		3300
3241	CCTCGGGACCTTGC CGGTTCGCGTTGAGCCTTATCATCTTTCAGTCCGATGGAGAGAAAGT	3300
3241		3300
3301	CATTGCTGGGGACGGAGACAGCTGCTTGTGGGGACAATTTTACCGGACTTCCGTGTC	3360
3301		3360
3301	CATGCTTGGGGACGGAGACAGCTGCTTGTGGGGAGTATCTTACCGGACTTCCGTGTC	3360
3301		3360
3361	GGCCGACTTGGTTCGGAGGTCCTCTTGGCCACGCTGATGGCTATACCTCCAGGGGGTG	3420
3361		3420
3361	GGCCGACTTGGTTCGGAGGTCCTCTTGGCCACGCTGATGGCTATACCTCCAGGGGGTG	3420
3361		3420
3421	GAGTCTTCTCGCCGCCCATCACTGCTTACGCCACGACAGCTGG	

QY 9361 GCCTGCCGACCCCGCTATTGCTCTTGGCCCTACTCTTCTTTTGTAGGGTAGGCT 9420
 DB |||||
 QY 9361 GCCTGCCGACCCCGCTATTGCTCTTGGCCCTACTCTTCTTTTGTAGGGTAGGCT 9420
 DB |||||
 QY 9421 TTTCTTACTCCCGCTCGGTAGAGCGGACACATAGTACACTCCATAGCTAACTGTCC 9480
 DB |||||
 QY 9421 TTTCTTACTCCCGCTCGGTAGAGCGGACACATAGTACACTCCATAGCTAACTGTCC 9480
 DB |||||
 QY 9481 CTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9540
 DB |||||
 QY 9481 CTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9540
 DB |||||
 QY 9541 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9588
 DB |||||
 QY 9541 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9588
 DB |||||
 RESULT 6
 LOCUS E07361
 DEFINITION GRNA of Hepatitis non-A non-B virus, HCV-J6.
 ACCESSION E07361
 VERSION E07361.1 GI:5708558
 KEYWORDS JP 1994121689-A/1.
 SOURCE unclassified
 ORGANISM unclassified.
 REFERENCE 1 (bases 1 to 9589)
 AUTHORS Okamoto.H. and Nakamura.T.
 TITLE NON-A NON-B HEPATITIS VIRUS GENE, POLYNUCLEOTIDE, POLYPEPTIDE,
 ANTIGEN AND ANTIBODY DETECTION SYSTEM
 JOURNAL Patent: JP 1994121689-A 1 06-MAY-1994;
 NAKAMURA TETSUO
 COMMENT OS Hepatitis non-A non-B virus
 PN JP 1994121689-A/1
 PD 06-MAY-1994
 PF 09-AUG-1991 JP 1991287402
 PI OKAMOTO HIROAKI, NAKAMURA TETSUO
 PC C12N15/51,A61K39/29,A61K39/395,A61K39/395,C07K13/00,C12P21/02,
 C12P21/08,
 PC G01N33/53,G01N33/576,G01N33/577//A61B10/00;
 CC strandedness: Single;
 CC topology: Linear;
 CC hypothetical: No;
 CC anti-sense: No; Location/Qualifiers
 FH key
 FT source 1..9589
 FT /organism="Hepatitis non-A non-B virus" FT
 /strains="HCV-J6".
 FEATURES
 source Location/Qualifiers
 1..9589
 /organism="unidentified"
 /mol_type="genomic RNA".
 /db_xref="taxon:32644"
 BASE COUNT 1968 a 2820 c 2635 g 2166 t
 ORIGIN
 Query Match 92.4%; Score 8970.4; DB 6; Length 9589;
 Best Local Similarity 96.0%; Pred. No. 0;
 Matches 9202; Conservative 0; Mismatches 386; Indels 0; Gaps 0;
 QY 1 ACCGGCCCTTAATAGGGGCGACACTCCGCATGATCACTCCCTGTGAGGAAGTACTGT 60
 DB |||||
 QY 1 ACCGGCCCTTAATAGGGGCGACACTCCGCATGATCACTCCCTGTGAGGAAGTACTGT 60
 DB |||||
 QY 61 CTTTCCAGCAGAAAGCGCTAGCCATGCGTTAGTATGAGTGTCTGACAGCTCCAGGCC 120
 DB |||||
 QY 61 CTTTCCAGCAGAAAGCGCTAGCCATGCGTTAGTATGAGTGTCTGACAGCTCCAGGCC 120
 DB |||||
 QY 121 CCCCCTCCGGGAGAGCATAGTGGTCTGCGGAACCGGTAGTACACCGGAATTGCGGG 180
 DB |||||

DB 121 CCCCCTCCGGGAGAGCATAGTGGTCTGCGGAACCGGTAGTACACCGGAATTGCGGG 180
 QY 181 AAGACTGGTCTCTTTCTTGGATAAACCCACTCTATGSCCGGCATTTTGGCGCTGCCCG 240
 DB |||||
 DB 181 AAGACTGGTCTCTTTCTTGGATAAACCCACTCTATGSCCGGCATTTTGGCGCTGCCCG 240
 QY 241 CAAGACTGCTAGCCGAGTAGCTTGGTTTGCAGAGGCTTTGTGTGATCTGCTGTAGGG 300
 DB |||||
 DB 241 CAAGACTGCTAGCCGAGTAGCTTGGTTTGCAGAGGCTTTGTGTGATCTGCTGTAGGG 300
 QY 301 TGTCTTCGAGTCCCGGGGAGGCTCTGCTAGACCGGTGCACCATGAGCACAATCTTAACC 360
 DB |||||
 DB 301 TGTCTTCGAGTCCCGGGGAGGCTCTGCTAGACCGGTGCACCATGAGCACAATCTTAACC 360
 QY 361 TCAAGAGAAAAACCAAAAGAAAAACCAAAACCGTCCGCCCAAGACGTTAAGTTTCCGGGCG 420
 DB |||||
 DB 361 TCAAGAGAAAAACCAAAAGAAAAACCAAAACCGTCCGCCCAAGACGTTAAGTTTCCGGGCG 420
 QY 421 CGGCCAGATCTGTGGCGAGTATATCTTGTTCGCGCGAGGGGCCAGGTGGGTGTGG 480
 DB |||||
 DB 421 CGGCCAGATCTGTGGCGAGTATATCTTGTTCGCGCGAGGGGCCAGGTGGGTGTGG 480
 QY 481 CGCGACAGGAAGACTTTCGAGGCGGTCCAGCCACGTCGGAAGCGCCAGCCATCCCTAA 540
 DB |||||
 DB 481 CGCGACAGGAAGACTTTCGAGGCGGTCCAGCCACGTCGGAAGCGCCAGCCATCCCTAA 540
 QY 541 AGATCGGCGCTCCACTGGCAAAATCTTGGGAAAAACCAAGATACCCCTGGCCCCCTATAC 600
 DB |||||
 DB 541 AGATCGGCGCTCCACTGGCAAAATCTTGGGAAAAACCAAGATACCCCTGGCCCCCTATAC 600
 QY 601 GAATGAGGACTCTGCGTGGCAGGATGGTCTCTGTCCTCCCGAGGTTTCCGCTCTCTT 660
 DB |||||
 DB 601 GAATGAGGACTCTGCGTGGCAGGATGGTCTCTGTCCTCCCGAGGTTTCCGCTCTCTT 660
 QY 661 GGGCCCCAATGACCCCGCGATAGTCTGCGCAACGTTGGGTAAAGTTCATCGATACCCCTAAC 720
 DB |||||
 DB 661 GGGCCCCAATGACCCCGCGATAGTCTGCGCAACGTTGGGTAAAGTTCATCGATACCCCTAAC 720
 QY 721 GTGCGCTTTTGGCGACTCATGCGGTGTACATCTCTGTCGCGGCGCCCGCTCGCGCGCG 780
 DB |||||
 DB 721 GTGCGCTTTTGGCGACTCATGCGGTGTACATCTCTGTCGCGGCGCCCGCTCGCGCGCG 780
 QY 781 CGCAGAGCTCTCGCGCATGCGGTGTAGAGTCTCTGAGGAGCGGGGTAAATTTTGCACAGG 840
 DB |||||
 DB 781 CGCAGAGCTCTCGCGCATGCGGTGTAGAGTCTCTGAGGAGCGGGGTAAATTTTGCACAGG 840
 QY 841 GAATTTACCGGTTGCTCTCTTTCTATCTTCTTGTGCGGCTGCTCTCTGCTGCTACCCAC 900
 DB |||||
 DB 841 GAATTTACCGGTTGCTCTCTTTCTATCTTCTTGTGCGGCTGCTCTCTGCTGCTACCCAC 900
 QY 901 CCGGCTCTCCGCTGCGCAAGTGAAGAACATCAGTACCGGTACATGCTGATTAACGACTG 960
 DB |||||
 DB 901 CCGGCTCTCCGCTGCGCAAGTGAAGAACATCAGTACCGGTACATGCTGATTAACGACTG 960
 QY 961 CACCAATGACAGCATTTACCTGGCAGCTCCAGGCTGCTGCTCTCCAGCTCCCGGGTCCGT 1020
 DB |||||
 DB 961 CACCAATGACAGCATTTACCTGGCAGCTCCAGGCTGCTGCTCTCCAGCTCCCGGGTCCGT 1020
 QY 1021 CCGCTGCGAGAAAGTGGGGAATGCACTCTCAGTCTGATACCGGTCTCACCGGAATGTGGC 1080
 DB |||||
 DB 1021 CCGCTGCGAGAAAGTGGGGAATGCACTCTCAGTCTGATACCGGTCTCACCGGAATGTGGC 1080
 QY 1081 CGTGACGAGCCCGCGGCGCTTACGAGGCTTACGAGCGCATTTGACATGTTGTAT 1140
 DB |||||
 DB 1081 CGTGACGAGCCCGCGGCGCTTACGAGGCTTACGAGCGCATTTGACATGTTGTAT 1140
 QY 1141 GTCCGCAAGCTCTGCTGCT 1200
 DB |||||
 DB 1141 GTCCGCAAGCTCTGCTGCT 1200
 QY 1201 AGCCCAATTTGTTTCTTCTCGCGGAGCAGCCACTGTTGTTGTTGTTGTTGTTGTTGTT 1260
 DB |||||
 DB 1201 AGCCCAATTTGTTTCTTCTCGCGGAGCAGCCACTGTTGTTGTTGTTGTTGTTGTTGTT 1260

Db	7141	GCCTTCGATACCATCGGAATATATGCTCCCAAGAGAGATTCCACACGCTTACCGGC	7200	Db	8221	GGTGATGGGGGCTTCTTATGATTCAGTACTCCCGCTCAGCGGTGAGTTCTCTT	8280
Qy	7201	CTGGGACCGGCTGATTACAAACCGCTTGTGGAACTCGTGAAGAGCCAGATTACCA	7260	Qy	8281	GAAGCATGGGCGGAAAGAACCCCTATGGGTTTTCTGATGATACCCGATGCTTTGA	8340
Db	7201	TTGGGACGCGCTGATTACAAACCGCTTGTGGAACTCGTGAAGAGCCAGATTACCA	7260	Db	8281	GAAGCATGGGCGGAAAGAACCCCTATGGGTTTTCTGATGATACCCGATGCTTTGA	8340
Qy	7261	ACCGGCACTGTTGGCGGCTGTCTCTCCCTCTCCTAGGAAACCCCGACGCTCCGCC	7320	Qy	8341	CTCAACCGTCACTGAGAGAGACATCAGGACTGAGGAGTCCATATATCGGGCTTCTCTT	8400
Db	7261	ACCGGCACTGTTGGCGGCTGTCTCTCCCTCTCCTAGGAAACCCCGACGCTCCGCC	7320	Db	8341	CTCAACCGTCACTGAGAGAGACATCAGGACTGAGGAGTCCATATATCGGGCTTCTCTT	8400
Qy	7321	AAGGAGCGCGGACAGTGGGCTTAAGTGAGGACTCAATAGGAGATGCCCTTCAACAGCT	7380	Qy	8401	GCCGAGGAGGCCACATCTGCCATACATCTCGCTAACTGAGAGACTTTAGTGGGAGGCC	8460
Db	7321	AAGGAGCGCGGACAGTGGGCTTAAGTGAGGACTCAATAGGAGATGCCCTTCAACAGCT	7380	Db	8401	GCCGAGGAGGCCACATCTGCCATACATCTCGCTAACTGAGAGACTTTAGTGGGAGGCC	8460
Qy	7381	GGCCATTAGTCTTTGGCCAGCCCCCCCCAGCGGCTTCAAGGCTTTTCCACGGGGC	7440	Qy	8461	TATGTTCAACAGCAAGGCGCAACCTCGGGGTACAGGCGTTCGCGCGCAGCGGGTGT	8520
Db	7381	GGCCATTAGTCTTTGGCCAGCCCCCCCCAGCGGCTTCAAGGCTTTTCCACGGGGC	7440	Db	8461	TATGTTCAACAGCAAGGCGCAACCTCGGGGTACAGGCGTTCGCGCGCAGCGGGTGT	8520
Qy	7441	GGCGCTGCCGATTCGCGGAGTCAAGCGCTCTGATGAGTTGGCCCTTTTCGAGACAGG	7500	Qy	8521	CACCACTAGCATGGGGAACACCATCACTGCTGAAAGCTTAGCGGCTTGTAAAGC	8580
Db	7441	GGCGCTGCCGATTCGCGGAGTCAAGCGCTCTGATGAGTTGGCCCTTTTCGAGACAGG	7500	Db	8521	TACCACTAGCATGGGGAACACCATCACTGCTGAAAGCTTAGCGGCTTGTAAAGC	8580
Qy	7501	TTCCATCTCTTCATGCCCTCCCTCGAGGGGAGCTTGAGATCCAGACTCTGAGCCTGA	7560	Qy	8581	TGCAAGGATTAATCGCGCCACAATGCTGATGCGCGGATGACTTGGTTGCTCATCAGA	8640
Db	7501	TTCCATCTCTTCATGCCCTCCCTCGAGGGGAGCTTGAGATCCAGACTCTGAGCCTGA	7560	Db	8581	TGCAAGGATTAATCGCGCCACAATGCTGATGCGCGGATGACTTGGTTGCTCATCAGA	8640
Qy	7561	GCAGGTAGAGCCCAACCCCCCCCCAGGGGGGTGGCAGCTCCCGGCTCGACTCGG	7620	Qy	8641	AAGCCAGGGGACCGAGGAGGAGCGAGGAACTTGAGAGCCTTCACGAGGCTATGACCAG	8700
Db	7561	GCAGGTAGAGCCCAACCCCCCCCCAGGGGGGTGGCAGCTCCCGGCTCGACTCGG	7620	Db	8641	AAGCCAGGGGACCGAGGAGGAGCGAGGAACTTGAGAGCCTTCACGAGGCTATGACCAG	8700
Qy	7621	GTCTGCTTACTTGTCTCGAGGAGGAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG	7680	Qy	8701	GTATTTCTGCCCCCTCTGCTGAGCCCCCGGAGCGGAGTATGATCTGGAGCTGATAACATC	8760
Db	7621	GTCTGCTTACTTGTCTCGAGGAGGAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG	7680	Db	8701	GTATTTCTGCCCCCTCTGCTGAGCCCCCGGAGCGGAGTATGATCTGGAGCTGATAACATC	8760
Qy	7681	CTGGACCGGGCTCTAATACTCTTGTAGTCCGAGGAGGAGGTTACCGATTAAACCC	7740	Qy	8761	TTGCTCTCAAAATGTTGTTGTTGCGCTGCGGCCCAAGGCCCGCGCAGATATCTAGAC	8820
Db	7681	CTGGACCGGGCTCTAATACTCTTGTAGTCCGAGGAGGAGGTTACCGATTAAACCC	7740	Db	8761	TTGCTCTCAAAATGTTGTTGTTGCGCTGCGGCCCAAGGCCCGCGCAGATATCTAGAC	8820
Qy	7741	CTTGAGCACTCCCTGTTGGGATATCAACAAGGTGTTACTGTACCAACAAGAGCGC	7800	Qy	8821	CAGAGACCTTACCACTCCAACTCGCCGGGCTGCTGCGGAAACAGTTAGACATCCCTCTG	8880
Db	7741	CTTGAGCACTCCCTGTTGGGATATCAACAAGGTGTTACTGTACCAACAAGAGCGC	7800	Db	8821	CAGAGACCTTACCACTCCAACTCGCCGGGCTGCTGCGGAAACAGTTAGACATCCCTCTG	8880
Qy	7801	CTCACTAAGGGCTAAAGGTAATTTGATAGATGCAAGTGTCTGACTCTCTACTACGA	7860	Qy	8881	CAATTTCTGCGTGGGAAACATCATCCAGTACGCCCGGAGCCATATGGGCTCGATGTCCT	8940
Db	7801	CTCACTAAGGGCTAAAGGTAATTTGATAGATGCAAGTGTCTGACTCTCTACTACGA	7860	Db	8881	CAATTTCTGCGTGGGAAACATCATCCAGTACGCCCGGAGCCATATGGGCTCGATGTCCT	8940
Qy	7861	CTCAGTCTTAAGGACATTAAGCTAGCGGCTTCAAGGTCAAGGCTCTCAACAT	7920	Qy	8941	GATGACACATCTTCTCTCAATTCATGGCCCAAGATATCTGAGCCAGAACTCAACTT	9000
Db	7861	CTCAGTCTTAAGGACATTAAGCTAGCGGCTTCAAGGTCAAGGCTCTCAACAT	7920	Db	8941	GATGACACATCTTCTCTCAATTCATGGCCCAAGATATCTGAGCCAGAACTCAACTT	9000
Qy	7921	GGAGGAGGCTGCGAGTTAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCT	7980	Qy	9001	TGAGATGTACGGAGCGGTGATCTCGGTGAGTCCCTTGGACCTCCAGCCATTAATTAAG	9060
Db	7921	GGAGGAGGCTGCGAGTTAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCT	7980	Db	9001	TGAGATGTACGGAGCGGTGATCTCGGTGAGTCCCTTGGACCTCCAGCCATTAATTAAG	9060
Qy	7981	TAGGAGGCTCGGAGCTTGTCCGGAGGCGGTTAACACATCAAGTCTCGTGGAGGA	8040	Qy	9061	GTATACATGGGCTTGACGCTTTTCTCTGACACATACACTCCCGGAGGCTGACCGGT	9120
Db	7981	TAGGAGGCTCGGAGCTTGTCCGGAGGCGGTTAACACATCAAGTCTCGTGGAGGA	8040	Db	9061	GTATACATGGGCTTGACGCTTTTCTCTGACACATACACTCCCGGAGGCTGACCGGT	9120
Qy	8041	CCTCTCGGAGGACTCAGAAACCAATTCACCAACCTTATGGCCCAAAATGAGGTGT	8100	Qy	9121	GGCTTCAGCCCTCAGAAACCTTGGGGCGCACCTTCAGAGCGTGGAGAGCGGGCAGC	9180
Db	8041	CCTCTCGGAGGACTCAGAAACCAATTCACCAACCTTATGGCCCAAAATGAGGTGT	8100	Db	9121	GGCTTCAGCCCTCAGAAACCTTGGGGCGCACCTTCAGAGCGTGGAGAGCGGGCAGC	9180
Qy	8101	CTGCGTGGAGCCCAAGGGGGGGAAGAGAGCTCGCTTATGCTTATCGCTGACCT	8160	Qy	9181	TGCACTGAGGCGTCCCTCATCTCCGTTGGGGAGAGCGGCGTTTGGGTCGATATCT	9240
Db	8101	CTGCGTGGAGCCCAAGGGGGGTAAGAAAGAGCTCGCTTATGCTTATCGCTGACCT	8160	Db	9181	TGCACTGAGGCGTCCCTCATCTCCGTTGGGGAGAGCGGCGTTTGGGTCGATATCT	9240
Qy	8161	CGGCGTCAAGGCTCTGAGAGAGTGGCCCTTTATGATGATTAACAAAGCTTCTCAGGC	8220	Qy	9241	CTTCAATTTGGGCGGTGAAGACCAAGCTCACTCTCATTTGCCGAGAGCGGCTCTCT	9300
Db	8161	CGGCGTCAAGGCTCTGAGAGAGTGGCCCTTTATGATGATTAACAAAGCTTCTCAGGC	8220	Db	9241	CTTCAATTTGGGCGGTGAAGACCAAGCTCACTCTCATTTGCCGAGAGCGGCTCTCT	9300
Qy	8221	GGTATGGGGGCTTCTTATGAGTTCCAGTATTTCCCGCTCAGCGGCTAGAGTTCTCTT	8280	Qy	9301	GGATTTATCCAGCTGGTTCACTGTCGGCGCGCGCGGCGGCGGCGGCGGCGGCGG	9360
Db	8221	GGTATGGGGGCTTCTTATGAGTTCCAGTATTTCCCGCTCAGCGGCTAGAGTTCTCTT	8280	Db	9301	GGATTTATCCAGCTGGTTCACTGTCGGCGCGCGCGGCGGCGGCGGCGGCGGCGG	9360

QY	4981	CAACACGCCCGTTTGCTGTGTCGCCAAGACCATTCTTGAGTTTGGAGGCGAGTTTTCAC	5040
DB	4981	CAACACACCTGGTTTGCTGTGTCGCCAAGACCATTCTTGAGTTTGGAGGCGAGTTTTCAC	5040
QY	5041	CGGCTCACACATAGATGCCACTTCTTCTCCCAACAAAGCAATCGGGGGAATTT	5100
DB	5041	CGGCTCACACATAGATGCCACTTCTTCTCCCAACAAAGCAATCGGGGGAATTT	5100
QY	5101	CGCATCTTAAACAGCTTACAGGCTACAGTGTGCGCTAGGGCCAAAGCCCCCGCTC	5160
DB	5101	CGCATCTTAAACAGCTTACAGGCTACAGTGTGCGCTAGGGCCAAAGCCCCCGCTC	5160
QY	5161	CTGGGACGTCAATGTGGAAGTGTGTTGACTCGACTCAAGCCACACTCGTGGGCCACACC	5220
DB	5161	CTGGGACGTCAATGTGGAAGTGTGTTGACTCGACTCAAGCCACACTCGTGGGCCACACC	5220
QY	5221	TCCTCTGACCGCTTGGGCTCTGTGTACCAACGAGGTCACCTCAACCTCTGTGACGAA	5280
DB	5221	TCCTCTGACCGCTTGGGCTCTGTGTACCAACGAGGTCACCTCAACCTCTGTGACGAA	5280
QY	5281	ATACATCGCCACTGTCATGCAAGCGACTTGTAGGTCATGACAGACACATGCGTCTTGGC	5340
DB	5281	ATACATCGCCACTGTCATGCAAGCGACTTGTAGGTCATGACAGACACATGCGTCTTGGC	5340
QY	5341	AGGGGAGTCTTGGCGGCGTGGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGATCAT	5400
DB	5341	AGGGGAGTCTTGGCGGCGTGGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGATCAT	5400
QY	5401	CGGCGCTTGACATTAACAGAGGAGCGCTGCTGCGCGGACCAAGAGGTCCTCTATGA	5460
DB	5401	CGGCGCTTGACATTAACAGAGGAGCGCTGCTGCGCGGACCAAGAGGTCCTCTATGA	5460
QY	5461	GGCTTTTGATGAGTGGAGGAATGTGCCCTGCGGCTGCGGCTGCTGCGGCTGCTGCG	5520
DB	5461	GGCTTTTGATGAGTGGAGGAATGTGCCCTGCGGCTGCGGCTGCTGCGGCTGCTGCG	5520
QY	5521	GATAGCCGAGATGCTGAAAGTCCAAGATCCAAGGCTTATTCAGCAAGCTTCCAAACAGC	5580
DB	5521	GATAGCCGAGATGCTGAAAGTCCAAGATCCAAGGCTTATTCAGCAAGCTTCCAAACAGC	5580
QY	5581	TCAGACATACACCCACTGTGAGGCTTCAATGAGGCTTCAATGAGGCTTCAATGAGGCTTCA	5640
DB	5581	TCAGACATACACCCACTGTGAGGCTTCAATGAGGCTTCAATGAGGCTTCAATGAGGCTTCA	5640
QY	5641	ACACATGTGGAATCTTATGAGGCTTCAATGAGGCTTCAATGAGGCTTCAATGAGGCTTCA	5700
DB	5641	ACACATGTGGAATCTTATGAGGCTTCAATGAGGCTTCAATGAGGCTTCAATGAGGCTTCA	5700
QY	5701	GAACCTGTGAGTGTTCATGATGGCTTCAAGTGGGCTTCAAGTGGGCTTCAAGTGGGCTTCA	5760
DB	5701	GAACCTGTGAGTGTTCATGATGGCTTCAAGTGGGCTTCAAGTGGGCTTCAAGTGGGCTTCA	5760
QY	5761	AAGCACCATATCTTCAACATTTTGGGGGCTGCTAGCATCCCAATGACACACC	5820
DB	5761	AAGCACCATATCTTCAACATTTTGGGGGCTGCTAGCATCCCAATGACACACC	5820
QY	5821	CGGGGGGCTACCGGCTTGTGTCAGTGGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGT	5880
DB	5821	CGGGGGGCTACCGGCTTGTGTCAGTGGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGT	5880
QY	5881	CTTAGGTAGGTGTAGTGGACATCTTGGGAGGCTTGGTGGGCTTGGTGGGCTTGGTGGGCTT	5940
DB	5881	CTTAGGTAGGTGTAGTGGACATCTTGGGAGGCTTGGTGGGCTTGGTGGGCTTGGTGGGCTT	5940
QY	5941	CGTGGATTCAGATCATGTCTGGGAGAGCCCTCCATGGAGGATGTGTCAACTGTCT	6000
DB	5941	CGTGGATTCAGATCATGTCTGGGAGAGCCCTCCATGGAGGATGTGTCAACTGTCT	6000
QY	6001	GCTCGAATTTCTGCTCGGGTGTCTGAGTGGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTT	6060
DB	6001	GCTCGAATTTCTGCTCGGGTGTCTGAGTGGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTT	6060

[illegible]

QY	601	GAATGAGGACTCGGCTGGGACGATGGCTCTGTCCCGCCGAGGTTCGCGTCCCTCTTG	660	QY	1681	CAACTCGTCAGATGTCCGAAACGATGTCCGCTGCGCAGTATCGAGGCTTCGCGGT	1740
DB	601	GAATGAGGACTCGGCTGGGACGATGGCTCTGTCCCGCCGAGGTTCGCGTCCCTCTTG	660	DB	1681	CAACTCGTCAGATGTCCGAAACGATGTCCGCTGCGCAGTATCGAGGCTTCGCGGT	1740
QY	661	GGGCCCCAATGATCCCGGATAGGTGCGCGAAGTGGGTAGGTGATGATACCTTAAC	720	QY	1741	GGGATGGGCGCTTCGAAATATGAGATAATGTCCACCAATCCAGAGGATATGAGACCTTA	1800
DB	661	GGGCCCCAATGATCCCGGATAGGTGCGCGAAGTGGGTAGGTGATGATACCTTAAC	720	DB	1741	GGGATGGGCGCTTCGAAATATGAGATAATGTCCACCAATCCAGAGGATATGAGACCTTA	1800
QY	721	GTGCGGCTTTGCGGACTCATGAGGTACATCCCTGTGCGGCGCCCGCTCGGCGGCGT	780	QY	1801	TTGCTGGCACTACCCACCAAGGACGTGTGGCGGTGTCTCCGCGAAGACTGTGTGTGGCCC	1860
DB	721	GTGCGGCTTTGCGGACTCATGAGGTACATCCCTGTGCGGCGCCCGCTCGGCGGCGT	780	DB	1801	TTGCTGGCACTACCCACCAAGGACGTGTGGCGGTGTCTCCGCGAAGACTGTGTGTGGCCC	1860
QY	781	CGCCAGAGCTCTCGCGCATGCGGTGAGAGTCTCTGGAGGACGGGTTAAATTTTGCAACAG	840	QY	1861	AGTGTACTGTCTTTCACCCCGAGCTGTAGTGGGACGACCGATAGACTTTGGAGCGCC	1920
DB	781	CGCCAGAGCTCTCGCGCATGCGGTGAGAGTCTCTGGAGGACGGGTTAAATTTTGCAACAG	840	DB	1861	AGTGTACTGTCTTTCACCCCGAGCTGTAGTGGGACGACCGATAGACTTTGGAGCGCC	1920
QY	841	GAATTTACCGGTTGCTCTTTTCTATCTTCTGCTGGGCGCTGCTGCTGCAATACCCAC	900	QY	1921	CACTTTACAGCTGGGGGAGATGAGACAGATGTCTTCTTATTTGAAACAGCACTCGACACCC	1980
DB	841	GAATTTACCGGTTGCTCTTTTCTATCTTCTGCTGGGCGCTGCTGCTGCAATACCCAC	900	DB	1921	CACTTTACAGCTGGGGGAGATGAGACAGATGTCTTCTTATTTGAAACAGCACTCGACACCC	1980
QY	901	CCCGGCTCTCGGCTGCGGAGTGAAGACATCAGTACCGGTACATGCTGACTAACGACTG	960	QY	1981	GCTGGGCTCATGCTGGCTGCGCTGCACTGGAATGAATCTCTTGGCTACACCAAGACTTTGCGG	2040
DB	901	CCCGGCTCTCGGCTGCGGAGTGAAGACATCAGTACCGGTACATGCTGACTAACGACTG	960	DB	1981	GCTGGGCTCATGCTGGCTGCGCTGCACTGGAATGAATCTCTTGGCTACACCAAGACTTTGCGG	2040
QY	961	CACCAATGACAGCAATTTACCTGGGAGTCCAGGCTGTCTCTCCAGTCTCCCGGGTGCCT	1020	QY	2041	CGCACCACTCTGCGCTGCTAGAGCTGACTTTCAAGCCGACGACGACCTGTGTGGCCCCAC	2100
DB	961	CACCAATGACAGCAATTTACCTGGGAGTCCAGGCTGTCTCTCCAGTCTCCCGGGTGCCT	1020	DB	2041	CGCACCACTCTGCGCTGCTAGAGCTGACTTTCAAGCCGACGACGACCTGTGTGGCCCCAC	2100
QY	1021	CCCGTGGGAAAGTGGGAAATCATCTAGTCTGATACCGGTCTACCGAATGTGGC	1080	QY	2101	GGAGTGTGTTAGGAGGATCTGTATACACTTACCTCAATGGGCTCTGGGCGCTGGGCT	2160
DB	1021	CCCGTGGGAAAGTGGGAAATCATCTAGTCTGATACCGGTCTACCGAATGTGGC	1080	DB	2101	GGAGTGTGTTAGGAGGATCTGTATACACTTACCTCAATGGGCTCTGGGCGCTGGGCT	2160
QY	1081	CGTGACGCGCCGCGGCTCTACGAGGCTTTGCGGACGACATCGACATGTTGTGAT	1140	QY	2161	CACGCCAAGGTGCTGATCGACTTACCCCTACAGGCTCTGGCAATTTACCCCTGCAACAGTTAA	2220
DB	1081	CGTGACGCGCCGCGGCTCTACGAGGCTTTGCGGACGACATCGACATGTTGTGAT	1140	DB	2161	CACGCCAAGGTGCTGATCGACTTACCCCTACAGGCTCTGGCAATTTACCCCTGCAACAGTTAA	2220
QY	1141	GTCCGCGCACGCTCTGCTCGCTCTTACGTGGGGACCTCTGCGGTGGGCTGATGCTGC	1200	QY	2221	CTATACCATCTTCAAAATTAAGATGTATGTGGAGGGGTTGAGCACAGGCTCAGGCTGC	2280
DB	1141	GTCCGCGCACGCTCTGCTCGCTCTTACGTGGGGACCTCTGCGGTGGGCTGATGCTGC	1200	DB	2221	CTATACCATCTTCAAAATTAAGATGTATGTGGAGGGGTTGAGCACAGGCTCAGGCTGC	2280
QY	1201	AGCCCAATGTTTCAATGTTCTGCGGACGACCACTGGTGTGCAAGCTGCAATTTGCTC	1260	QY	2281	ATGCAATTTTCACTCGTGGGGATCGTTGCAACTTTGGAGGACAGAGACAGAACTCACTGTC	2340
DB	1201	AGCCCAATGTTTCAATGTTCTGCGGACGACCACTGGTGTGCAAGCTGCAATTTGCTC	1260	DB	2281	ATGCAATTTTCACTCGTGGGGATCGTTGCAACTTTGGAGGACAGAGACAGAACTCACTGTC	2340
QY	1261	CATCTACCTGTTACCATCACTGACACCGCATGGGATGGGATGATGATGAACTGGTC	1320	QY	2341	TCCTTTGTTGCACTCCACCAACGGAATGGCCCAATTTTACCTTCTTACTCGGACCTGCC	2400
DB	1261	CATCTACCTGTTACCATCACTGACACCGCATGGGATGGGATGATGATGAACTGGTC	1320	DB	2341	TCCTTTGTTGCACTCCACCAACGGAATGGCCCAATTTTACCTTCTTACTCGGACCTGCC	2400
QY	1321	GCCACAGGCTTACCATGATCTTGGCGTACGCGATGCGGTGCTCCCGAGGTCAATTTAGACAT	1380	QY	2401	CGCCTTGTGCACTGTGCTTCTCCACCTCCACCAAAACATCGTGGAGCTGCAATTCATGTA	2460
DB	1321	GCCACAGGCTTACCATGATCTTGGCGTACGCGATGCGGTGCTCCCGAGGTCAATTTAGACAT	1380	DB	2401	CGCCTTGTGCACTGTGCTTCTCCACCTCCACCAAAACATCGTGGAGCTGCAATTCATGTA	2460
QY	1381	CATTAGCGGGCTCATTTGGGCGCTCATGTTGCGGTGGGCTACTTCTCTATGAGGGAGC	1440	QY	2461	TGGGCTATCACCTGCGCTCACAATAATACATCGTCCGATGGAGTGGGTATCTTTATT	2520
DB	1381	CATTAGCGGGCTCATTTGGGCGCTCATGTTGCGGTGGGCTACTTCTCTATGAGGGAGC	1440	DB	2461	TGGGCTATCACCTGCGCTCACAATAATACATCGTCCGATGGAGTGGGTATCTTTATT	2520
QY	1441	GTGGGCAAGTGTGTTGATCTTCTGTTGGCGCGGGTGGACGCGGACCCATAC	1500	QY	2521	CTGTCTTTAGCGGACGCGGCTTGGCGCTTCTTATGATGCTCATCTTGTGTGGGCCA	2580
DB	1441	GTGGGCAAGTGTGTTGATCTTCTGTTGGCGCGGGTGGACGCGGACCCATAC	1500	DB	2521	CTGTCTTTAGCGGACGCGGCTTGGCGCTTCTTATGATGCTCATCTTGTGTGGGCCA	2580
QY	1501	TGTTGGGGGTTCTGCGCGCAGACCAACCGGCGCTCACCAGCTTATTTTGACATGGGCCC	1560	QY	2581	GGCCGAAGGACGACCTAGAGAAAGTTGGTCTTCTTGCACGCTGCCGAGGCGAGTGCATA	2640
DB	1501	TGTTGGGGGTTCTGCGCGCAGACCAACCGGCGCTCACCAGCTTATTTTGACATGGGCCC	1560	DB	2581	GGCCGAAGGACGACCTAGAGAAAGTTGGTCTTCTTGCACGCTGCCGAGGCGAGTGCATA	2640
QY	1561	CAGCGAATAATCCAGTCTGTTAAACCAATGAGGAGCTGGGACATCAACCGGACCGGCT	1620	QY	2641	TGGCTTCTTATATTTTGTCTATCTTTTCTGTTGGCTGCTTGTATCATCAAGGGTGGGTAGT	2700
DB	1561	CAGCGAATAATCCAGTCTGTTAAACCAATGAGGAGCTGGGACATCAACCGGACCGGCT	1620	DB	2641	TGGCTTCTTATATTTTGTCTATCTTTTCTGTTGGCTGCTTGTATCATCAAGGGTGGGTAGT	2700
QY	1621	GAATGCAATGACTCTTGGACACCGGCTTATGCGGCTCTGTTCTACACCCACAGCTT	1680	QY	2701	CCCTTTAGCTTACCTATTTCCCTCAGTGGGCTGTGGTCTTTTAGCTTACTGTCTCTAGCATT	2760
DB	1621	GAATGCAATGACTCTTGGACACCGGCTTATGCGGCTCTGTTCTACACCCACAGCTT	1680	DB	2701	CCCTTTAGCTTACCTATTTCCCTCAGTGGGCTGTGGTCTTTTAGCTTACTGTCTCTAGCATT	2760
				QY	2761	GCCCCAACAGGCTTATGCTTTATGACGATCTGTGCAATGCGCCAGATAGGAGCGGCTCTGCT	2820

ORGANISM

Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepadnavirus.

REFERENCE

1 (bases 1 to 9589)
Okamoto, H., Okada, S., Sugiyama, Y., Kurai, K., Izuka, H., Machida, A.,
Miyakawa, Y., and Miyumi, M.
Nucleotide sequence of the genomic RNA of hepatitis C virus
isolated from a human carrier: comparison with reported isolates
for conserved and divergent regions
J. Gen. Virol. 72 (Pt 11), 2697-2704 (1991)

JOURNAL

MEDLINE

PUBMED

1658196

REFERENCE

2 (sites)

AUTHORS

Han, J. H. and Houghton, M.

TITLE

Group specific sequences and conserved secondary structures at the
3' end of HCV genome and its implication for viral replication

Nucleic Acids Res. 20 (13), 3520 (1992)

JOURNAL

MEDLINE

PUBMED

1321416

REFERENCE

3 (sites)

AUTHORS

Hotta, H., Doi, H., Hayashi, T., Purwanta, M., Soemarto, W.,
Mizokami, M., Ohba, K. and Homma, M.

TITLE

Analysis of the core and E envelope region sequences of a novel
variant of hepatitis C virus obtained in Indonesia

Arch. Virol. 136 (1-2), 53-62 (1994)

JOURNAL

MEDLINE

PUBMED

9427090

COMMENT

7545932

These data kindly submitted in computer readable form by: Hiroaki
Okamoto

Immunology Division

Jichi Medical School

Kawachi-gun

Tochigi-ken 329-04

Japan

Phone: 0285-44-2111 x3334

Fax: 0285-44-1557.

Location/Qualifiers

1. .9589

/organism="Hepatitis C virus"

/mol_type="genomic RNA"

/isolate="HC-J6"

/db_xref="taxon:11103"

341. .9442

/codon_start=1

/product="polyprotein"

/protein_id="BA00792.1"

/db_xref="GI:21651"

/translation="MSTNPKPQKTKRNNRRPQDKPFGGQIVGGVYLLPRRGRPL

GYRATKTSERSQPRGRQPIPKDRRTCKSGKFGYEPWLYGNEGLGAWGLISPRG

SRPSGPNDRHRNRNKGVIDLTCGFADLMGYIPVVGAPLGGVARALAHGVRLV

GVNFATGNLPGCSRSIFLLALSLCITTPVSAAEVKNISTGVYVNDCTNDSITWLOA

AVLHVPQVCPCKEKGNTSRCHLPVSPNVAVQPGALTOGLRTHIDMVVNSATLSALY

VDLCCGVMLAAQMTIVSPQHHWFVQDCNSIYPGTITGHRMAMDMNMNSPTATMIL

AYAMRVPIVDIICGAGHGMFGLAYFSMOGAWKVVVILLAAQVDAQTHVGGST

AHNRATLTCMSGLARQKQILQINTGNSHNINRNLNDSLHTGFLASLFTHSFNSS

GCPBMSACRSIEAFVGMGALQYEDNVNTPEDMRPYCHWYPRGCGVSVASVCGVP

YCFPSGVVLTDLRGLAPTYTWGENETDVLNSTRPQGSWFCTWNSGTGKTC

GAPPCRIRADFNASMDLCLPDCFRKHEDTITKCGSGFWLTPRCLIDIPRLWHYPC

TYNTYIFKRMVGVGVEHRLTAACNFTGRDCNEDRDSQLSPILLSTTEWAILPCT

YSDLPALSTGLLHLQNTIVDQPMGLSPALTYIWRWVYLLFLLLADARVCACLW

MLILGQAALAEKLVLWHAASAACNGFLYFVIFVAAWYIKGRVPLATYSITGLW

SPGLLLALPOQAVYADASVHGOIGAILLVLITLFTLTPGKYLILSRFLWLCYLLTL

AEAMVQENAPMQVGGRGDIIMAWAIFCPGVVEDITKWLAVLGLPALLKGLTRVP

YFVRHALRMTVMRHLAGRYVOMLLALRGRTGTIYIIDLTPMSDWAANGLRDLA

VAPEPIIFSPMEKKVIVMGAETAACGDIHLGLPVSARGREVLGLPADGYTSKWSLL

APITAYAGLGLLGTIVVMTGRDKTEQAGEIQVLSVTQSTFLGTITVGLWTVYHG

AKNTLAGSRGPOTWYSAEDGLVWSPSPCKSLPCTCGAVDLVLTNRADVIPA

RNRGKRALSLSPRLSTLKGSGGVLCPGRHAGVGFRAAVCSRGVAKSIDFIPVET

LDIVRSPFTSDNSTPRAVPQYQVYGLHAPTGSCKSTKVPVAYAAQYKVLVLPVS

AATLGFSLSKAHGNIENRIGRVITGTAPITSTYSGKFLADGGCAGGAYDIIICD

ECHAVDSTILIGVLDQAEATAGVRLTVLATATPPGVSVTTPHPNIEEVLQSGEIP

FYGRAIPLSYIKGGRHLLFCHSKKKCDLBAALRGMLNAVYYRGLDVSIVPTQGDV

VVAVDALMTGFTGDFDSVIDCNVAVTQVDFSLDPTFTIITQTVPQDAVRSRRGR

FEATURES

source

CDS

QY	1	ACCCGCCCTAATAGGGCGACACATCCGCCATGAATCACTCCCTGTGAGGAATCTACTGT	60
DB	1	ACCCGCCCTAATAGGGCGACACATCCGCCATGAATCACTCCCTGTGAGGAATCTACTGT	60
QY	61	CTTCACGAGAAAGCGCTAGCCATGGCGTTAGTATAGTGTCTGTACAGCCTCCAGGCC	120
DB	61	CTTCACGAGAAAGCGCTAGCCATGGCGTTAGTATAGTGTCTGTACAGCCTCCAGGCC	120
QY	121	CCCCCTCCCGGAGAGCCATAGTGTCTGCGGAACCGGTGAGTACACCGGAATTCGCCGG	180
DB	121	CCCCCTCCCGGAGAGCCATAGTGTCTGCGGAACCGGTGAGTACACCGGAATTCGCCGG	180
QY	181	AAGACTGGGTCTCTTTCTTGGATAAACCCACTCTATGCCCGGTCAATTTGGCGGTGCCCG	240
DB	181	AAGACTGGGTCTCTTTCTTGGATAAACCCACTCTATGCCCGGTCAATTTGGCGGTGCCCG	240
QY	241	CAAGACTGCTAGCGAGTAGCGTTGGGTTCGGAAGCGCTTGTGTACTGTCTGATAGGG	300
DB	241	CAAGACTGCTAGCGAGTAGCGTTGGGTTCGGAAGCGCTTGTGTACTGTCTGATAGGG	300
QY	301	TGCTTTCAGGTGCCCCCGGAGGTCTGTGTAGACCGGTGACCATGACACAAATCTCTAAACC	360
DB	301	TGCTTTCAGGTGCCCCCGGAGGTCTGTGTAGACCGGTGACCATGACACAAATCTCTAAACC	360
QY	361	TCAAAGAAAAACAAAAAGAAAAACCAACCGTCGCCCAACAAGACGTAAAGTTTCGGGGCGG	420
DB	361	TCAAAGAAAAACAAAAAGAAAAACCAACCGTCGCCCAACAAGACGTAAAGTTTCGGGGCGG	420
QY	421	CGGCCAGATCGTTGGCGGAGTATACTTGTTCGCCGGAGGGGCCCGAGTTGGGTGTGGG	480
DB	421	CGGCCAGATCGTTGGCGGAGTATACTTGTTCGCCGGAGGGGCCCGAGTTGGGTGTGGG	480
QY	481	CGCGAAGAAAGACTTCGAGCGGTCCACACCGTGGAAAGGGCCAGCCATCCCTTAA	540
DB	481	CGCGAAGAAAGACTTCGAGCGGTCCACACCGTGGAAAGGGCCAGCCATCCCTTAA	540
QY	541	AGATCGGCGTCTCACTGGCAAACTCTGGGGAAAAACACAGGATACCCCTGGCCCCCTATACGG	600
DB	541	AGATCGGCGTCTCACTGGCAAACTCTGGGGAAAAACACAGGATACCCCTGGCCCCCTATACGG	600

BASE COUNT 1970 a 2820 c 2633 g 2166 t
ORIGIN

Query Match 92.4%; Score 8973.6; DB 14; Length 9589;
Best Local Similarity 96.0%; Pred. No. 0;
Matches 9204; Conservative 0; Mismatches 384; Indels 0; Gaps 0;

QY	7681	CTGGACCGGGCTCTAATACCTCTTGTAGTCCCGAAGAGGAGTACCGATTAAACCC	7740	8761	TTGCTCTCAAAATGTGTCTGTGGCGCTGGGCCCAAGGCCCGCAGATACCTGAC	8820
DB	7681	CTGGACCGGGCTCTAATACCTCTTGTAGTCCCGAAGAGGAGTACCGATTAAACCC	7740	8761	TTGCTCTCAAAATGTGTCTGTGGCGCTGGGCCCAAGGCCCGCAGATACCTGAC	8820
QY	7741	CTTGAGCAACTCCCTCTGTGCGATATCACAAAGGTGTACTGTACCAACAAGAGCGC	7800	8821	CAGAGACCTTACCACTCCAAATCGCCGGCTGCTGGGAAACAGTTAGACACTCCCTGT	8880
DB	7741	CTTGAGCAACTCCCTCTGTGCGATATCACAAAGGTGTACTGTACCAACAAGAGCGC	7800	8821	CAGAGACCTTACCACTCCAAATCGCCGGCTGCTGGGAAACAGTTAGACACTCCCTGT	8880
QY	7801	CTCACTAAGGGCTAAAGGTAACTTTTGTATAGGATGCAAGTCTGACTCTCTACGGA	7860	8881	CAATTCATGGCTGGGAAACATCATCCAGTACGCCCGCCAGCATATGGGCTCGATGCTCT	8940
DB	7801	CTCACTAAGGGCTAAAGGTAACTTTTGTATAGGATGCAAGTCTGACTCTCTACGGA	7860	8881	CAATTCATGGCTGGGAAACATCATCCAGTACGCCCGCCAGCATATGGGCTCGATGCTCT	8940
QY	7861	CTCAGTCTTAAAGGACATTAAGCTAGCGGCTCCAAAGGTCCAGGCTCTCTACCAT	7920	8941	GATGACACACTTCTTCTCCATCTCATGGCTCAAGACACGCTGGACCAAGCTCAACTT	9000
DB	7861	CTCAGTCTTAAAGGACATTAAGCTAGCGGCTCCAAAGGTCCAGGCTCTCTACCAT	7920	8941	GATGACACACTTCTTCTCCATCTCATGGCTCAAGACACGCTGGACCAAGCTCAACTT	9000
QY	7921	GGAGGAGGCTTGCAGTTAAACCCACCCCAATCTGCAAGATCTAAATATGGGTTTGGGC	7980	9001	TCAGATGACGAGCGGTGTACTCCGTGAGTCCCTGGACCTCCAGCTATATTAATCAAG	9060
DB	7921	GGAGGAGGCTTGCAGTTAAACCCACCCCAATCTGCAAGATCTAAATATGGGTTTGGGC	7980	9001	TCAGATGACGAGCGGTGTACTCCGTGAGTCCCTGGACCTCCAGCTATATTAATCAAG	9060
QY	7981	TAAGGAGGCTCCGAGCTTGTCCGGAGGCGGTTTACCAATCAAGTCCGTGGGAGGA	8040	9061	GTTCATATGGCTTGACGCTTTTCTTGACACATACACTCCCCCAAGCTGACACGGGT	9120
DB	7981	TAAGGAGGCTCCGAGCTTGTCCGGAGGCGGTTTACCAATCAAGTCCGTGGGAGGA	8040	9061	GTTCATATGGCTTGACGCTTTTCTTGACACATACACTCCCCCAAGCTGACACGGGT	9120
QY	8041	CCTCTGGAGGACTCAGAAACACCAATTTCCACACCAATTAATGCGGTAAGGAGGA	8100	9121	GGCTCAGCGCTCAGAAACTTGGGGCGCCACCCCTCAGAGCGTGGAGCGGGCAGC	9180
DB	8041	CCTCTGGAGGACTCAGAAACACCAATTTCCACACCAATTAATGCGGTAAGGAGGA	8100	9121	GGCTCAGCGCTCAGAAACTTGGGGCGCCACCCCTCAGAGCGTGGAGCGGGCAGC	9180
QY	8101	CTCGTGGAGGCTCAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	8160	9181	TCGAGTCAGGGCTCCCTCATCTCCGTGGGGGAGAGCGGCGTTTGGGTCGATATCT	9240
DB	8101	CTCGTGGAGGCTCAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	8160	9181	TCGAGTCAGGGCTCCCTCATCTCCGTGGGGGAGAGCGGCGTTTGGGTCGATATCT	9240
QY	8161	CGGCGTCAGGGCTCTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	8220	9241	CTTCAATTCGGCGGTGAAGCAAGCTCAAACTCACTCATTTGCCGGAAGCGGCTCTCT	9300
DB	8161	CGGCGTCAGGGCTCTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	8220	9241	CTTCAATTCGGCGGTGAAGCAAGCTCAAACTCACTCATTTGCCGGAAGCGGCTCTCT	9300
QY	8221	GGTGTGGGGCTCTTATGAGTTCAGTATTCCTCCGCTCAGCGGTGAGGTTCTCT	8280	9301	GGATTTATCAGCTGTTTCAACCGTCCGCGCGCGCGCGCGCGCGCATTTATCACAGCGT	9360
DB	8221	GGTGTGGGGCTCTTATGAGTTCAGTATTCCTCCGCTCAGCGGTGAGGTTCTCT	8280	9301	GGATTTATCAGCTGTTTCAACCGTCCGCGCGCGCGCGCGCGCATTTATCACAGCGT	9360
QY	8281	GAAAGCATGGGCGGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	8340	9361	GGTGTCCGAGCGGCTTATGCTCTTGGCGCTTACTCTTCTTGTAGGGGTAGGCGCT	9420
DB	8281	GAAAGCATGGGCGGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	8340	9361	GGTGTCCGAGCGGCTTATGCTCTTGGCGCTTACTCTTCTTGTAGGGGTAGGCGCT	9420
QY	8341	CTCAACCGTCACTGAGAGAGACATCAGGAGTGGAGGAGTCCATATATCGGGCTT	8400	9421	TTTCTTACTCTCCGCTCGGTGAGCGGCACACATTAAGCTACACTCCATAGCTAACCTGTC	9480
DB	8341	CTCAACCGTCACTGAGAGAGACATCAGGAGTGGAGGAGTCCATATATCGGGCTT	8400	9421	TTTCTTACTCTCCGCTCGGTGAGCGGCACACATTAAGCTACACTCCATAGCTAACCTGTC	9480
QY	8401	GCCGAGGAGGCGCACATCGCTACACTCGCTAACTGAGAGACTTTTACGTGGAGGCG	8460	9481	CTTT	9540
DB	8401	GCCGAGGAGGCGCACATCGCTACACTCGCTAACTGAGAGACTTTTACGTGGAGGCG	8460	9481	CTTT	9540
QY	8461	TATGTTCAACAGCAAGGCGCAAACTCGCGGTACAGCGGTTCGCGCCAGCGGGTGT	8520	9541	TT	9600
DB	8461	TATGTTCAACAGCAAGGCGCAAACTCGCGGTACAGCGGTTCGCGCCAGCGGGTGT	8520	9541	TT	9600
QY	8521	CACCACTAGCATGGGGAACCATCATCTGCTAGGTGAAGGCTTACGGGCTTGAAGC	8580	9601	TACTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	9660
DB	8521	CACCACTAGCATGGGGAACCATCATCTGCTAGGTGAAGGCTTACGGGCTTGAAGC	8580	9601	TACTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	9660
QY	8581	TGCAGGATTAATCGCGCCCAATGCTGTTATGGGGGATGACTTGGTGTCTCTCAGA	8640	9661	GAGCCGATGACTGCAGAGAGTCCGTAACCTGCTCTCTGAGATCATGT 9711	
DB	8581	TGCAGGATTAATCGCGCCCAATGCTGTTATGGGGGATGACTTGGTGTCTCTCAGA	8640	9661	GAGCCGATGACTGCAGAGAGTCCGTAACCTGCTCTCTGAGATCATGT 9711	
QY	8641	AAGCAGGGGACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	8700			
DB	8641	AAGCAGGGGACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	8700			
QY	8701	GTATTCGCCCCCTCTCTGTTGACCCCGGAGCGGAGGATGATCTGAGCTGATAATC	8760			
DB	8701	GTATTCGCCCCCTCTCTGTTGACCCCGGAGCGGAGGATGATCTGAGCTGATAATC	8760			

RESULT 5
HPCPOLP
LOCUS
DEFINITION
Hepatitis C virus genomic RNA for polyprotein, complete cds.
D00944
VERSION
D00944.1
KEYWORDS
polyprotein.
Hepatitis C virus

2161	CACGCCAAGGTGCGCTGATCGACTACCCCTTACAGGCTCTGGCAATTACCCCTGACACAGTTAA	2220	CTATACCATCTTTCAAAATAAGAGATGTATGTGGAGGGGTTGAGCACAGGCTCACGGCTGCG	2280
2221		2281		
2221	CTATACCAATCTTTCAAAATAAGAGATGTATGTGGAGGGGTTGAGCACAGGCTCACGGCTGCG	2280		
2281	ATGCAATTTCACTCGTGGGGATCGTTGCAACTTGGAGGACAGAGACAGAAGTCAACTGTCT	2340		
2281	ATGCAATTTTCACTCGTGGGGATCGTTGCAACTTGGAGGACAGAGACAGAAGTCAACTGTCT	2340		
2341	TCCCTTTGTGTGACTCCACACAGGAATGGGCCAATTTTACCTTGTCTTCTTACTCGGACTCGCC	2400		
2341	TCCCTTTGTGTGACTCCACACAGGAATGGGCCAATTTTACCTTGTCTTCTTACTCGGACTCGCC	2400		
2401	CGCCTTGTGCACTGGTCTTCTCCACTCCACCAAAACATCGTGGACGTACAAATTCATGTAT	2460		
2401	CGCCTTGTGCACTGGTCTTCTCCACTCCACCAAAACATCGTGGACGTACAAATTCATGTAT	2460		
2461	TGGCCTATPACCTGCCCTCACAAAATAATCATCGTCCGATGGGAGTGGGTAATACTCTTAAT	2520		
2461	TGGCCTATPACCTGCCCTCACAAAATAATCATCGTCCGATGGGAGTGGGTAATACTCTTAAT	2520		
2521	CCTGCTCTTAGGGAGCGCAGGGTTTGGCCCTGCTTATGGATGCTCATCTTGTGGGGCA	2580		
2521	CCTGCTCTTAGGGAGCGCAGGGTTTGGCCCTGCTTATGGATGCTCATCTTGTGGGGCA	2580		
2581	GGCCGAGCAGCAGCTAGAGAAGCTGGTTCATCTTGCACGCTCGAGCGCAGCTAGCTGCAAT	2640		
2581	GGCCGAGCAGCAGCTAGAGAAGCTGGTTCATCTTGCACGCTCGAGCGCAGCTAGCTGCAAT	2640		
2641	TGGCTTCCATATATTTTGTTCATCTTTTTCGTGGCTGCTTTTACATCAAGGGTGGGTAGT	2700		
2641	TGGCTTCCATATATTTTGTTCATCTTTTTCGTGGCTGCTTTTACATCAAGGGTGGGTAGT	2700		
2701	CCCCTTAGTAGTACTTATTCCTTCACTGGGCTGTGGTCTTTAGGCTTACTGCTCCTAGCAAT	2760		
2701	CCCCTTAGTAGTACTTATTCCTTCACTGGGCTGTGGTCTTTAGGCTTACTGCTCCTAGCAAT	2760		
2761	GCCCAACAGGCTTATGCTTATGAGCATCTGTGCATGCCAGCCAGATAGGAGCGGCTCTGCT	2820		
2761	GCCCAACAGGCTTATGCTTATGAGCATCTGTGCATGCCAGCCAGATAGGAGCGGCTCTGCT	2820		
2821	GCTAATGATCACTCTCTTTTACTCTCACCCCGGGTATAAGACCCCTTCTCAGCCGGTTTT	2880		
2821	GCTAATGATCACTCTCTTTTACTCTCACCCCGGGTATAAGACCCCTTCTCAGCCGGTTTT	2880		
2881	GTGTGGTGTGCTATCTTCTGACCCCTGGGGGAAGCTATGGTCCAGGAGTGGGCACCAACC	2940		
2881	GTGTGGTGTGCTATCTTCTGACCCCTGGGGGAAGCTATGGTCCAGGAGTGGGCACCAACC	2940		
2941	TATCGAGTGGCGGTGGCGGTGATGGCATCATATGGGCCGTGCGCCATATCTACCCAGG	3000		
2941	TATCGAGTGGCGGTGGCGGTGATGGCATCATATGGGCCGTGCGCCATATCTACCCAGG	3000		
3001	TGTGGTGTGTGACATATAACAAAGTGGCTCTTGGCGGTGCTTGGGCTGCTTACTCCTAAA	3060		
3001	TGTGGTGTGTGACATATAACAAAGTGGCTCTTGGCGGTGCTTGGGCTGCTTACTCCTAAA	3060		
3061	AGGTGCTTTTGAAGCGCGTGGCGGTGATGGCATCATATGGGCCGTGCGCCATATCTACCCAGG	3120		
3061	AGGTGCTTTTGAAGCGCGTGGCGGTGATGGCATCATATGGGCCGTGCGCCATATCTACCCAGG	3120		
3121	CATGGCAGGCAATCTCGGGGGGCGAGTACGTCCAGATGGCGTACTAGCCCTTGGCAG	3180		
3121	CATGGCAGGCAATCTCGGGGGGCGAGTACGTCCAGATGGCGTACTAGCCCTTGGCAG	3180		
3181	GTGACTGGCACTTACATCTATGACCACTCAACCCCTATGTCCGATTTGGGCTGCTAGTGG	3240		
3181	GTGACTGGCACTTACATCTATGACCACTCAACCCCTATGTCCGATTTGGGCTGCTAGTGG	3240		
3241	CCTGGGACCTGGCGGTGCGGTTGAGCCTTATCATCTTCAGTCCGATGGAGAGAAAGT	3300		
3241	CCTGGGACCTGGCGGTGCGGTTGAGCCTTATCATCTTCAGTCCGATGGAGAGAAAGT	3300		

source	1..9711 /organism="Hepatitis C virus" /mol_type="genomic RNA" /strain="HC-J6CH" /db_xref="taxon:11103" /clone="pu6ct" /notes="infectious cDNA genotype: 2a" 341..9442 CDS
	/codon_start=1 /product="polyprotein" /protein_id="AA01178.1" /db_xref="GI:6010580" /translation="MSTNPKPQRKTRNTRRNPQDKVFGGGQIVGGVLLPRRGRPL GVRAKTKTSRSPRRRPIPKDRSTGKSGKPGYPMPLYNEGLGWAGLLSPRG SRPSGNDPRHSRNVKVIDPLTCGFADLMGYI PVGAPLGGVAAALAHGVRVLED GVNPAITLPGCSFSLIALLSCITTPVSAABVKNIISGYVWINDCTNDSITWQLOA AVLHVPCCVCEKGNASQWIPVSNVAORPQALTOGLRTHIDMWVMSATLCSALY VGLDCGVMLAQMFTVSPQHFVFLQYFSGMOGAWAKVVVILLIAGVDARTHTVGGSA AOTTRGLTSLFDMGRPKIQLVNTGNSWHINRTALNCNDSLHTGFTASIPYTHSFNS GCPERSACRSIEAFVGVGALQYEDVNTNPEMRPYCWHPROCGVVSAAKTVCPVP YCFPSFVVGTTRDLAGPTTYWGENETDVFLLNSTRPFLGFWFGCTWNSSGYTKC GAPPCRTADFNASTDLCTPCFRKHPDTTYLKGSGPWLTPRLCIDIPYRLWHVPC TVNYTTFKIRMYVGVGVEHLRTAACNFTGRDRCNLEDRSQSLPLHSTTEWAILPCS YSDLPALSTGLLHLHONI VDVQPMYGLSPALTKYI VRWEAWYILLFLLADARVCACLM MLTLGQAAEALFKVILHAASACNGFLYFVI FVFAAWYIKGRVVPLATVSLTGLM SFSLLIALLPOQAYADASHVGOIGAAILVIMTLFTLTPTGYKTLISRLFWLWLYLLTL GEAWQEPAPMVGVRDGI I WAVAIFYPGVVFDITKMLLAVIGPAYLLKGLATRPV YFVRAHALLRMTMARHLAGRYYQWALLAGVGTGTYYIDHLPMSDWAASGLRDLA VAVEPII FSPMEKRTVWGAETAACGDIILHGLFVSARLQREVLLGPDAGYTSKWSLL APITAAVQOTRGLGTVVSMTRQDKTEQAGEIVLSVTQSPFGTSISGLVWTVHG AGNKTLAGSGPVQMYSSAEGDLVGMPSPPGPKSLEPCTCGAVIDLYLVRNADVI PA RRGDRKGRLLSRPRLSTLKGSGGPVLCPRGHAVGVFRAAVCSRVAKSIDIPVET LDLIVTSPTSDNSTPAPVPTTQVGLHAPITGSGSKSTKVPVAIAQGIKVLVNSV AATLFGALYSKAGHNPINRTGRTVTGTAPITSTYTGKFLADGGCAGAYDIIICD ECHAVSTYTLIGTVDQAEATAGRLTVLATATPPGVTTPHNTPEEVALGOEBIP FYGAATPLSYIKGGRHILFCHSKKCDLAAALRGMLNSAYVYRGLDVSVP TQGDIP VVVATDALMTGTPDVSVDVNVAVTVVDDFSLDPTFTITTOI PQDQAVSRORGR TGRGLGIYRYVSTGERASGMPDVLCSYDAGAAYELTPTSETVRLRAYVNTPL VQCQDHLFEAVFTGTLTHIDHFLSQTSQGENPAYLTAQATVCARAKAPPPSDV MKCLRLHNVGTEDCPICPGGSLRDVMDVCTITLTDPKNWLTSKLPFKMPGLPFVSCQ GVLAAYAYCGLTCVCIIIGRINQRAVAVAPDKVLEYAFEDEMECASRAALIEEGQ RIAEMLKSLQGLLOQASQAOIQPTVOASMPKQVFWAKHWNFTISGQIYLAGIST LPCNPVASMWAFSAALTSPSLTSLTILNLGGLASOIAPAGATGFVVSGLVGA VGSIGLGLKVLVDLAGVAGISALVAFKIMSKEKPSMEDVNLPGIISPGALVVG ICAAIILRRHVGPGEAGVQMNRLIAFASRGNHVAPTHYVTVESDASORVTLGLSLIT SLRLRLHNVGTEDCPICPGGSLRDVMDVCTITLTDPKNWLTSKLPFKMPGLPFVSCQ KGYGVWAGTGMITTRPCGANISGNVRLGSMRITGPKTMNIWOQTFPINCYTEQC VPKAPNFKVAIWRVAASEYAVTQHSYHYITGLTDNLKVPCLSPSEFFFSWBDG QIHRFAPTPKPPRDEVSFCVGLNSFVVSQLPDPEPDDVDMLNLTDPSSHITASTA ARLRAGSPSEASSASQASLSLATCTTHGKAYDVMDVNDANLFWGDVTRIESG KVVLDSLDPMVEERDLEPSISEYMLPKRFPALPAPWARPDYAPPLVESWKRDPY QVATVACALPPRPKPTTPPPRRRTVGLSEDSIGDALQOLAITSFGQPPPSGDLG TGAGADSGSTQPDDELALSETGISIMPPLEGELGDPDLEPEQVFPQPPQGGVAAP GSDSGWSTCEEDSDVVCMSYSMTGALITPCSPEEKLPINLPSNLSRLRYHNKYV CTTTKASIAKAVTPODMOVLDSYVDSVLKDIKLAASKVTARLLTMEERACOLTPHS ARSKYGEKAVRSLSGRVNHISKVWKLLEDSETPIPTTIMAKNEVFCVDTKGGK KAARLIYVPLGVRCVCEKALYDITQKLPQAVMSAGSYFOYSPQAFVFLFKAWAEKK DPMGFSYTRFASCTVTERDITEESIYRACSLPEEAHTAIHSLTRLYVGGPMFNSK GOTCGYVRCRSGVLTITMGNTITCYKALAAKAAAGIITAPTMVCGDDIDLVISSEQ TEEDERLRAFTAMTRYSPAPGDPPEYDLELITSCSNVSVLALGPQGRRYLYLTR DPTTPPARAAWETVRHSPVNSWLNIIQVAPTIWABMVLMTHPESILMAQDTLDQNL FEMIVAGISVPLDPAIIEKHLGLDFAFSLHTYTPHELTRVSAALRKLGAAPLRANKS RARAVRASLISRGRAAVCGRYLFWNAVKTKJLTLPLPEARLLDLSWFTVAGGGDI YHSVSRAPRLLFLGLLFLFVGLFLLPAR"
	1987 a 2853 c 2668 g 2203 t
BASE COUNT	100.0%; Score 9711; DB 14; Length 9711; Query Match Best local Similarity 100.0%; Pred. No. 0; Matches 9711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
ORIGIN	

7921 GGAGGAGGTTGCCAGTTAAACCCACCCCAATCTGCAAGATCTAAATATGGGTTGGGC 7980
7981 TAAGAGGTCGCGAGCTTGTCCGGAGGGCGGTTAAACCAATCAAGTCCGTGTGAAGA 8040
7981 TAAGAGGTCGCGAGCTTGTCCGGAGGGCGGTTAAACCAATCAAGTCCGTGTGAAGA 8040
8041 CTTCTGGAGGACTCAGAAACACCAATTCACCAACATATGCGCCAAAATAGAGTGT 8100
8041 CTTCTGGAGGACTCAGAAACACCAATTCACCAACATATGCGCCAAAATAGAGTGT 8100
8101 CTGCGTGGACCCACCAAGGGGGCAAGAGAGCTCGCTTATCGTTTACCTGACCT 8160
8101 CTGCGTGGACCCACCAAGGGGGCAAGAGAGCTCGCTTATCGTTTACCTGACCT 8160
8161 CGCGCTCAGGCTCTGCCAGAGATGGCCCTTTATGACATTTACCAAAAATCTTCTCAGC 8220
8161 CGCGCTCAGGCTCTGCCAGAGATGGCCCTTTATGACATTTACCAAAAATCTTCTCAGC 8220
8221 GGTGATGGGGGCTTCTTATGGAATTCAGTATTTCCCGCTCAGCGGTAGAGTTCTCT 8280
8221 GGTGATGGGGGCTTCTTATGGAATTCAGTATTTCCCGCTCAGCGGTAGAGTTCTCT 8280
8281 GAAAGCATGGCGGAAAGAAAGAGATGAGGAGTCCATATATCGGGGCTGCTCCTT 8340
8281 GAAAGCATGGCGGAAAGAAAGAGATGAGGAGTCCATATATCGGGGCTGCTCCTT 8340
8341 CTCAACCGTCACTGAGAGAGATCAGGAGTCCATATATCGGGGCTGCTCCTT 8400
8341 CTCAACCGTCACTGAGAGAGATCAGGAGTCCATATATCGGGGCTGCTCCTT 8400
8401 GCCGAGGAGGCCACACATCGCTAATCGTAGAGACTTTAGTGGGAGGGCC 8460
8401 GCCGAGGAGGCCACACATCGCTAATCGTAGAGACTTTAGTGGGAGGGCC 8460
8461 TATGTTCAACAGAGGCGAAACCTGCGGCTACAGCGGTTGCGCGCCAGCGGGTGCT 8520
8461 TATGTTCAACAGAGGCGAAACCTGCGGCTACAGCGGTTGCGCGCCAGCGGGTGCT 8520
8521 CACCACTAGCATGGGGAACACATCATGCTACGTGAAGCTTAGCGGCTGTAAAGC 8580
8521 CACCACTAGCATGGGGAACACATCATGCTACGTGAAGCTTAGCGGCTGTAAAGC 8580
8581 TGACGGGATTAATCGCGCCCAATGCTGGTATCGCGGATGACTTGGTTGTCTCAGA 8640
8581 TGACGGGATTAATCGCGCCCAATGCTGGTATCGCGGATGACTTGGTTGTCTCAGA 8640
8641 AAGCAGGGGACCGAGGAGGACGAGCGGAACTTGAGAGCTTTACCGAGGCTATGACCA 8700
8641 AAGCAGGGGACCGAGGAGGACGAGCGGAACTTGAGAGCTTTACCGAGGCTATGACCA 8700
8701 GTATTCTGCCCTCTCTGAGTACCCCGGAGCGGAGTATGCTGAGCTATGACATC 8760
8701 GTATTCTGCCCTCTCTGAGTACCCCGGAGCGGAGTATGCTGAGCTATGACATC 8760
8761 TTGCTCTCAATGCTCTGAGTACCCCGGAGCGGAGTATGCTGAGCTATGACATC 8820
8761 TTGCTCTCAATGCTCTGAGTACCCCGGAGCGGAGTATGCTGAGCTATGACATC 8820
8821 CAGAGACCTTACACTTCAATCGCGGCTGCTGGGAAACAGTTAGACATCCCGTGT 8880
8821 CAGAGACCTTACACTTCAATCGCGGCTGCTGGGAAACAGTTAGACATCCCGTGT 8880
8881 CAATTATCGGCGGAAACATCATCGTACCGCCCGGAGCGGAGTATGCTGAGCTATGACATC 8940
8881 CAATTATCGGCGGAAACATCATCGTACCGCCCGGAGCGGAGTATGCTGAGCTATGACATC 8940
8941 GATGACACACTTCTCTCAATCTCTATGCTCAAGACAGCTGAGCAAGCACTCAACT 9000
8941 GATGACACACTTCTCTCAATCTCTATGCTCAAGACAGCTGAGCAAGCACTCAACT 9000
9001 TGAGATGTAAGGAGCGGTACTCGGTAGTCCCTTGGACCTCCCGAGCTATTAATGAAG 9060
9001 TGAGATGTAAGGAGCGGTACTCGGTAGTCCCTTGGACCTCCCGAGCTATTAATGAAG 9060

9061 GTTATCATGGGCTGACGGCTTTTCTCTGCACATACACTCCCGACGAACTGACAGGCT 9120
9061 GTTATCATGGGCTGACGGCTTTTCTCTGCACATACACTCCCGACGAACTGACAGGCT 9120
9121 GCGTTACGCTTACAGAAACCTTGGGGCGCACCCCTCAGAGCGTGGAGAGCGGGCAGC 9180
9121 GCGTTACGCTTACAGAAACCTTGGGGCGCACCCCTCAGAGCGTGGAGAGCGGGCAGC 9180
9181 TGCACTCAGGGGCTCCCTCATCTCCGTCGGGGGAGAGCGGCTTTCGGTTCGATATCT 9240
9181 TGCACTCAGGGGCTCCCTCATCTCCGTCGGGGGAGAGCGGCTTTCGGTTCGATATCT 9240
9241 CTTCAATTTGGGCGGTGAAGACCAAGCTCAATTCCTCATTTGCCGGAAGCGGCTCCT 9300
9241 CTTCAATTTGGGCGGTGAAGACCAAGCTCAATTCCTCATTTGCCGGAAGCGGCTCCT 9300
9301 GGATTTATCCAGCTGGTTACCGTCCGGCGCGGGGGCGACATTTATCACAGCGTGTCC 9360
9301 GGATTTATCCAGCTGGTTACCGTCCGGCGCGGGGGCGACATTTATCACAGCGTGTCC 9360
9361 GCGTCCCGGACCCCGCTTATGCTCTTTGGGCTACTCTCTCTTTTGTAGGGGTAGGCT 9420
9361 GCGTCCCGGACCCCGCTTATGCTCTTTGGGCTACTCTCTCTTTTGTAGGGGTAGGCT 9420
9421 TTTCTCTACTCCCGCTCGGTAGAGCGGCACACATTTAGCTTACACTCCATAGTAACTGTCC 9480
9421 TTTCTCTACTCCCGCTCGGTAGAGCGGCACACATTTAGCTTACACTCCATAGTAACTGTCC 9480
9481 CTTT 9540
9481 CTTT 9540
9541 TT 9600
9541 TT 9600
9601 TACTTTCTCTCTGCTGCTCATCTTAGCCCTAGCTCAGCGCTAGCTGTAAGAGTCCCT 9660
9601 TACTTTCTCTCTGCTGCTCATCTTAGCCCTAGCTCAGCGCTAGCTGTAAGAGTCCCT 9660
9661 GAGCGCATGACTGCAGAGAGTCCGTAACCTGCTCTCTCTGAGATCATGT 9711
9661 GAGCGCATGACTGCAGAGAGTCCGTAACCTGCTCTCTCTGAGATCATGT 9711

RESULT 4

AF177036
LOCUS
DEFINITION
Hepatitis C virus clone pJ6CF, complete genome.
ACCESSION
AF177036
VERSION
AF177036.1
KEYWORDS
GI:6010579
SOURCE
Hepatitis C virus
ORGANISM
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
REFERENCE
1 (bases 1 to 9711)
Yanagi, M., Purcell, R.H., Emerson, S.U. and Bukh, J.
Hepatitis C virus: an infectious molecular clone of a second major genotype (2a) and lack of viability of intertypic 1a and 2a chimera
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
2 (bases 1 to 9711)
Bukh, J.
Direct Submission
Submitted (11-AUG-1999) Hepatitis Viruses Section, Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Building 7, Room 201, 7-Center, Bethesda, MD 20892-0740, USA
JOURNAL
TITLE
JOURNAL
FEATURES
Virology 262 (1), 250-263 (1999)

Qy	5761	AAGCACCACATCTCTTCAACATTTTGGGGGCTGGCTAGCATCCCAATTCACACACC	5820	6841	CCCTGAAACCCGACACAGAGTATTGATGTCCATTAACAGATCCATCTCATATCACGGC	6900	
Db	5761	AAGCACCACATCTCTTCAACATTTTGGGGGCTGGCTAGCATCCCAATTCACACACC	5820	Qy	6901	GGAGACTCAGCGCGCGCTTTAGCGCGGGGTACCCCATCCGAGGCAAGCTCTCTCGGC	6960
Qy	5821	CGCGGGGCGCACTGGCTTCGTTGCAAGTGCCTTAGTGGAGCTGCCGTAGCGCATAGG	5880	Db	6901	GGAGACTCAGCGCGCGCTTTAGCGCGGGGTACCCCATCCGAGGCAAGCTCTCTCGGC	6960
Db	5821	CGCGGGGCGCACTGGCTTCGTTGCAAGTGCCTTAGTGGAGCTGCCGTAGCGCATAGG	5880	Qy	6961	GAGCCAGCTATCGGCAACCATCGCTCGAGCCACTGCTGACACCCACGCAAGCCCTATGA	7020
Qy	5881	CTTAGTGAAGTCTAGTGACATCTCGGAGAGTATGGTCGGGCAATTTCCGGGGCTCT	5940	Db	6961	GAGCCAGCTATCGGCAACCATCGCTCGAGCCACTGCTGACACCCACGCAAGCCCTATGA	7020
Db	5881	CTTAGTGAAGTCTAGTGACATCTCGGAGAGTATGGTCGGGCAATTTCCGGGGCTCT	5940	Qy	7021	TGTGACATGTTGGATGCTTAACCTGTTTATGGGGGCGATGTGACTCGATAGAGTCTGG	7080
Qy	5941	CGTCGCAATCAAGATCATGTCGCGAGAGCCCTCCATGGAGATGTCGTCAACTTGGT	6000	Db	7021	TGTGACATGTTGGATGCTTAACCTGTTTATGGGGGCGATGTGACTCGATAGAGTCTGG	7080
Db	5941	CGTCGCAATCAAGATCATGTCGCGAGAGCCCTCCATGGAGATGTCGTCAACTTGGT	6000	Qy	7081	GTCCAAAGTGTGTTCTGGACTCTCTCGACCAATGTCGAAAGAGAGGACCTTGA	7140
Qy	6001	GCTTGAATTCGTCTCCGGTGTCTTGGTAGTGGAGTCAATCGCGGCCATTTCTGGG	6060	Db	7081	GTCCAAAGTGTGTTCTGGACTCTCTCGACCAATGTCGAAAGAGAGGACCTTGA	7140
Db	6001	GCTTGAATTCGTCTCCGGTGTCTTGGTAGTGGAGTCAATCGCGGCCATTTCTGGG	6060	Qy	7141	GCCTTCGATACCATCAGATATCATGCTCCCAAGAGAGGTTCCCAACAGCTTTACCGGC	7200
Qy	6061	CCGACAGTGGAGCCGGGAGCGCCGTCGAATGATGAATAGACTCAATTCGCTTTGC	6120	Db	7141	GCCTTCGATACCATCAGATATCATGCTCCCAAGAGAGGTTCCCAACAGCTTTACCGGC	7200
Db	6061	CCGACAGTGGAGCCGGGAGCGCCGTCGAATGATGAATAGACTCAATTCGCTTTGC	6120	Qy	7201	CTGGGCAAGGCTGATTACAAACCCGCTTCTGGAATCGTGGAAAGGCCAGATTACCA	7260
Qy	6121	TTCCAGAGGAATCAGCTCGCCGCCACCCACTAGTACGAGTCCGATCGTCGACGG	6180	Db	7201	CTGGGCAAGGCTGATTACAAACCCGCTTCTGGAATCGTGGAAAGGCCAGATTACCA	7260
Db	6121	TTCCAGAGGAATCAGCTCGCCGCCACCCACTAGTACGAGTCCGATCGTCGACGG	6180	Qy	7261	ACCGGCACTGTTGGCGGTGCTCTCCCTCTCTAGGAAACCCCGACGCTCTCCCC	7320
Qy	6181	TGTGACCCAACTACTTGGCTCCCTTACATAACCCAGCTCTCTCAGAGACTCCACAATG	6240	Db	7261	ACCGGCACTGTTGGCGGTGCTCTCCCTCTCTAGGAAACCCCGACGCTCTCCCC	7320
Db	6181	TGTGACCCAACTACTTGGCTCCCTTACATAACCCAGCTCTCTCAGAGACTCCACAATG	6240	Qy	7321	AAGGAGGCGCGGACAGTGGGCTTAAGTGAAGACTCCATAGGAGATGCTTCAACAGCT	7380
Qy	6241	GATTTAGGAGTCCCACTCCATCGCGGCTCGTGGCTCCGCGATGTTGGGACTG	6300	Db	7321	AAGGAGGCGCGGACAGTGGGCTTAAGTGAAGACTCCATAGGAGATGCTTCAACAGCT	7380
Db	6241	GATTTAGGAGTCCCACTCCATCGCGGCTCGTGGCTCCGCGATGTTGGGACTG	6300	Qy	7381	GGCCATTAAGTCTTTGGCGAGCCCGCCCAAGCGGCGATTCAGGCTTTTCAACGGGGC	7440
Qy	6301	GGTTTGACCACTTCAACAGACTTTAAATTTGGCTGACCTCCAAATTTATCCCAAGAT	6360	Db	7381	GGCCATTAAGTCTTTGGCGAGCCCGCCCAAGCGGCGATTCAGGCTTTTCAACGGGGC	7440
Db	6301	GGTTTGACCACTTCAACAGACTTTAAATTTGGCTGACCTCCAAATTTATCCCAAGAT	6360	Qy	7441	GGGCGCTGCGGATTCGCGCAGTCCAGACGCTCTCTGATGATGTTGGCCCTTTTCGAGACAGG	7500
Qy	6361	GCCCGGCTCCCTTTGCTCTCTGTAAGAGGGGTACAAGGGCTGTGGGCGCGCACTGG	6420	Db	7441	GGGCGCTGCGGATTCGCGCAGTCCAGACGCTCTCTGATGATGTTGGCCCTTTTCGAGACAGG	7500
Db	6361	GCCCGGCTCCCTTTGCTCTCTGTAAGAGGGGTACAAGGGCTGTGGGCGCGCACTGG	6420	Qy	7501	TTCCATCTCTTCATGCCCCCTCGAGGGGAGCTTGAGATCCAGACTGAGAGCTGA	7560
Qy	6421	CATCATGACACAGGCTGCTTGGCGGCGCAATATCTTGGCAATCTCGCTTGGGCTC	6480	Db	7501	TTCCATCTCTTCATGCCCCCTCGAGGGGAGCTTGAGATCCAGACTGAGAGCTGA	7560
Db	6421	CATCATGACACAGGCTGCTTGGCGGCGCAATATCTTGGCAATCTCGCTTGGGCTC	6480	Qy	7561	GCAGGTAGAGCCCAACCCCGGCGGCTTGGCAGCTCCCGGCTCGGACTCGGG	7620
Qy	6481	CATGAGATCAAGGGGCTTAAGACTTGAATATCTGGCAGGGGACCTTTCTCTATCAA	6540	Db	7561	GCAGGTAGAGCCCAACCCCGGCGGCTTGGCAGCTCCCGGCTCGGACTCGGG	7620
Db	6481	CATGAGATCAAGGGGCTTAAGACTTGAATATCTGGCAGGGGACCTTTCTCTATCAA	6540	Qy	7621	GTCTCTGTCTTACTTGTCTCGAGGAGACGACTCCGTCTGTCTGTCTTCCATGTCTACTC	7680
Qy	6541	TTGTTACACGAGGCGCAGTGGTGGGAAACCCCGCGCAACTTTAAGGTGCGCATCTG	6600	Db	7621	GTCTCTGTCTTACTTGTCTCGAGGAGACGACTCCGTCTGTCTGTCTTCCATGTCTACTC	7680
Db	6541	TTGTTACACGAGGCGCAGTGGTGGGAAACCCCGCGCAACTTTAAGGTGCGCATCTG	6600	Qy	7681	CTGGACCGGGCTCTAATACTCTTGTAGTCCGAGAGGAGAGTTCACGATTAAACCC	7740
Qy	6601	GAGGTCGCGGCTCAGAGTACGCGAGGTGACGACGCGGTATACCACTACATAAC	6660	Db	7681	CTGGACCGGGCTCTAATACTCTTGTAGTCCGAGAGGAGAGTTCACGATTAAACCC	7740
Db	6601	GAGGTCGCGGCTCAGAGTACGCGAGGTGACGACGCGGTATACCACTACATAAC	6660	Qy	7741	CTTGAGCAACTCCCTGTTGGGATATCAACAGAGTGTACTGTACCAACAAGAGCGC	7800
Qy	6661	AGGACTCACCACTGATTAACCTTGAAGTCCCTGCGCAACTCCCTCCCGAGTTCTTTTC	6720	Db	7741	CTTGAGCAACTCCCTGTTGGGATATCAACAGAGTGTACTGTACCAACAAGAGCGC	7800
Db	6661	AGGACTCACCACTGATTAACCTTGAAGTCCCTGCGCAACTCCCTCCCGAGTTCTTTTC	6720	Qy	7801	CTCAGTAAAGGCTTAAAGGTAACTTTTGTATAGATGCAAGTCTCGACTCTTACTACGA	7860
Qy	6721	CTGGGTGAGCGAGTGCAGATCCATAGGTTTGGCCCGCACCGAGCGGTTTTCGGGA	6780	Db	7801	CTCAGTAAAGGCTTAAAGGTAACTTTTGTATAGATGCAAGTCTCGACTCTTACTACGA	7860
Db	6721	CTGGGTGAGCGAGTGCAGATCCATAGGTTTGGCCCGCACCGAGCGGTTTTCGGGA	6780	Qy	7861	CTCAGTCTTAAAGGCACTTAAGCTAGCGGCTCCAAAGGTCCACGCAAGGCTCTTACCAT	7920
Qy	6781	TGAGTCTCGTCTCGTGGCTTAATTTCAATTTGTGTCGGGTCCGAGCTTCTTTCGGA	6840	Db	7861	CTCAGTCTTAAAGGCACTTAAGCTAGCGGCTCCAAAGGTCCACGCAAGGCTCTTACCAT	7920
Db	6781	TGAGTCTCGTCTCGTGGCTTAATTTCAATTTGTGTCGGGTCCGAGCTTCTTTCGGA	6840	Qy	7921	GGAGGAGGCTTGGCAGTTAAACCCCACTTCTGCAAGATCTTAATATGGTTTGGGC	7980
Qy	6841	CCCTGAACCCGACACAGAGCTATTGATGTCCATGCTTAACAGATCCATCTCATATCACGGC	6900				

3541 CA CAGTCACTACCTCTCTCCGGAACATCCATCTCGGGGGTTTGTGGACTGTCTACCA 3600
 3601 TGAGCTGGCAACAAGACTCTGGCCGGCTCAGCGGGTCCGGTCAACGACATGACTCCAG 3660
 3601 TGAGCTGGCAACAAGACTCTGGCCGGCTCAGCGGGTCCGGTCAACGACATGACTCCAG 3660
 3661 TGCTGAGGGGACTTATAGTGGGTCGACGCCCCCTGGGACTAAATCTTTGGAGCCGTG 3720
 3661 TGCTGAGGGGACTTATAGTGGGTCGACGCCCCCTGGGACTAAATCTTTGGAGCCGTG 3720
 3721 CAGCTGTGAGCGGTGACCTGTACCTGTGTCAGCGGAAACGCTGATGTATCCGGCTCG 3780
 3721 CAGCTGTGAGCGGTGACCTGTACCTGTGTCAGCGGAAACGCTGATGTATCCGGCTCG 3780
 3781 AAGACGCGGGCAAAACGGGAGCGGTACTCTCCCGGAGACCTCTTTCCACCTTGAAGGG 3840
 3781 AAGACGCGGGCAAAACGGGAGCGGTACTCTCCCGGAGACCTCTTTCCACCTTGAAGGG 3840
 3841 GTCTCTCAGGAGGCCCGGTGTATGCCCGAGGGCCACGCTGTCCGAGTCTTCCGGGCAGC 3900
 3841 GTCTCTCAGGAGGCCCGGTGTATGCCCGAGGGCCACGCTGTCCGAGTCTTCCGGGCAGC 3900
 3901 TGTGTGCTCTCGGGGGTGGCTAAGTCCATAGATTTTCACTCCCGTTGAGACACTCGACAT 3960
 3901 TGTGTGCTCTCGGGGGTGGCTAAGTCCATAGATTTTCACTCCCGTTGAGACACTCGACAT 3960
 3961 CGTCAACGCGTCCCGACCTTTAGTGAACAACAGCAGCACACCTGTGTCGCCAGACCTA 4020
 3961 CGTCAACGCGTCCCGACCTTTAGTGAACAACAGCAGCACACCTGTGTCGCCAGACCTA 4020
 4021 TCAGGTCCGGTACTTGCATGCCCGGACTGGAGTGGAAAGAGCACCAAAAGTTCTGTGCG 4080
 4021 TCAGGTCCGGTACTTGCATGCCCGGACTGGAGTGGAAAGAGCACCAAAAGTTCTGTGCG 4080
 4081 ATATGTCTCAGGGGTATAAGTGTAGTGTCTTAATCCCTCAGTGGCTGCCACCTGGG 4140
 4081 ATATGTCTCAGGGGTATAAGTGTAGTGTCTTAATCCCTCAGTGGCTGCCACCTGGG 4140
 4141 GTTTGGGGGTACTTGTCTAAGGCAATGSCATCAATCCCAATTTAGGACTGGAGTCA 4200
 4141 GTTTGGGGGTACTTGTCTAAGGCAATGSCATCAATCCCAATTTAGGACTGGAGTCA 4200
 4201 GACTGTGACGACCGGGGCGCCATCAGTACTCCACATATGGCAATTCCTCGCCGATGG 4260
 4201 GACTGTGACGACCGGGGCGCCATCAGTACTCCACATATGGCAATTCCTCGCCGATGG 4260
 4261 GGGCTGTGGGGCGGGCGCTACGACATCATATATGTGATGAATGCCATGCCGTGGACTC 4320
 4261 GGGCTGTGGGGCGGGCGCTACGACATCATATATGTGATGAATGCCATGCCGTGGACTC 4320
 4321 TACCACTATCTTGGCATCGGACAGTCTTGTATCAAGCAGACAGCTGGGGTCAAGT 4380
 4321 TACCACTATCTTGGCATCGGACAGTCTTGTATCAAGCAGACAGCTGGGGTCAAGT 4380
 4381 AACTGTGCTGGTACAGTCTACGCCCCCTGGGTGAGTGAACACCCCGCCCAACATAGA 4440
 4381 AACTGTGCTGGTACAGTCTACGCCCCCTGGGTGAGTGAACACCCCGCCCAACATAGA 4440
 4441 GGAGTGGCCCTTGGGCGAGGCGGAGATCCCTTCTATGGAGGGCGAATCCCTGTTC 4500
 4441 GGAGTGGCCCTTGGGCGAGGCGGAGATCCCTTCTATGGAGGGCGAATCCCTGTTC 4500
 4501 TTACATCAAGGAGGAGACATCTGATCTTCTGCCATTCAGAGAAAGTGTGACGCT 4560
 4501 TTACATCAAGGAGGAGACATCTGATCTTCTGCCATTCAGAGAAAGTGTGACGCT 4560
 4561 CGCGGCGCCCTTGGGGTATGGGCTTGAACCTCAGTGGCATACTACAGAGGGTGGAGCT 4620
 4561 CGCGGCGCCCTTGGGGTATGGGCTTGAACCTCAGTGGCATACTACAGAGGGTGGAGCT 4620
 4621 CTCCTGAATACCACTCAGGAGACGCTAGTGGTGTGCGCACCGACGCTCATGACAGG 4680
 4621 CTCCTGAATACCACTCAGGAGACGCTAGTGGTGTGCGCACCGACGCTCATGACAGG 4680

4681 GTATACTGGGAGCTTTGACTCCGTGATCGACTGCAACGCTAGCGGTCACTCAAGTTGTAGA 4740
 4681 GTATACTGGGAGCTTTGACTCCGTGATCGACTGCAACGCTAGCGGTCACTCAAGTTGTAGA 4740
 4741 CTTTCACTTTAGACCCCACTTACCAATTAACCAACAGATTTCCCTCAAGACGCTGTCTC 4800
 4741 CTTTCACTTTAGACCCCACTTACCAATTAACCAACAGATTTCCCTCAAGACGCTGTCTC 4800
 4801 AGCTAGCAGCGCCGGGTGCGACCGGTAGGGGAAGACTGGGCATTTATAGTATGTTTC 4860
 4801 AGCTAGCAGCGCCGGGTGCGACCGGTAGGGGAAGACTGGGCATTTATAGTATGTTTC 4860
 4861 CACTGTGAGGAGCGCTCAGGAATCTTTGACAGTGTAGTGTCTCTGTAGTGTCTAGACGC 4920
 4861 CACTGTGAGGAGCGCTCAGGAATCTTTGACAGTGTAGTGTCTCTGTAGTGTCTAGACGC 4920
 4921 AGGGCCCGCATGTATGAGCTCACACCATCGAGACCAACCGTCAAGGCTCAGGGCGTATTT 4980
 4921 AGGGCCCGCATGTATGAGCTCACACCATCGAGACCAACCGTCAAGGCTCAGGGCGTATTT 4980
 4981 CAAACGCGCGGTTTGCCTGTGTGCAAGACCATCTTGAGTTTGGGAGGCGATTTTTCAC 5040
 4981 CAAACGCGCGGTTTGCCTGTGTGCAAGACCATCTTGAGTTTGGGAGGCGATTTTTCAC 5040
 5041 CGGCTCTCACACATAGATGCCCTCTCTTCCCAACAAAGCAATCGGGGAAATTT 5100
 5041 CGGCTCTCACACATAGATGCCCTCTCTTCCCAACAAAGCAATCGGGGAAATTT 5100
 5101 CGCATATTAAACAGCTTACAGGCTACAGTGTGCGCTAGGGCCAAAGCCCCCCCCCGTC 5160
 5101 CGCATATTAAACAGCTTACAGGCTACAGTGTGCGCTAGGGCCAAAGCCCCCCCCCGTC 5160
 5161 CTGGGACGCTCATGTGGAAAGTGTGACTGCTGACTCAGGCCCCACACTGCTGGGCCCCACAC 5220
 5161 CTGGGACGCTCATGTGGAAAGTGTGACTGCTGACTCAGGCCCCACACTGCTGGGCCCCACAC 5220
 5221 TCTCTGTGCTACCGTTGGGCTCTGTTTACCAACGAGTCACTCAGATCCGCTGACCA 5280
 5221 TCTCTGTGCTACCGTTGGGCTCTGTTTACCAACGAGTCACTCAGATCCGCTGACCA 5280
 5281 ATACATCGCCACCTGCTGATGCAAGCGGCTTTGAGTGTATGACGACATGCGTCTTGGC 5340
 5281 ATACATCGCCACCTGCTGATGCAAGCGGCTTTGAGTGTATGACGACATGCGTCTTGGC 5340
 5341 AGGGGAGTCTTGGGCGCGCTGCGCGCTATTTGCTGGCGACCGGGTGTGTTGATCAT 5400
 5341 AGGGGAGTCTTGGGCGCGCTGCGCGCTATTTGCTGGCGACCGGGTGTGTTGATCAT 5400
 5401 CGGCGCTTGCACATTAACACGAGCGCTGTTGCGCGGACAGGAGTCTCTATGA 5460
 5401 CGGCGCTTGCACATTAACACGAGCGCTGTTGCGCGGACAGGAGTCTCTATGA 5460
 5461 GGCCTTTGATGAGATGAGGAGTGTGCTCTAGGGCGGCTCTCAATTGAAGAGGGGACGG 5520
 5461 GGCCTTTGATGAGATGAGGAGTGTGCTCTAGGGCGGCTCTCAATTGAAGAGGGGACGG 5520
 5521 GATAGCCAGATGTGAAGTCAAAGTCTTATGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGG 5580
 5521 GATAGCCAGATGTGAAGTCAAAGTCTTATGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGG 5580
 5581 TCAAGACATACCAACCTCTGAGGCTTCAATGCGCCCAAGGTAGAACAAATTTCTGGGCCAA 5640
 5581 TCAAGACATACCAACCTCTGAGGCTTCAATGCGCCCAAGGTAGAACAAATTTCTGGGCCAA 5640
 5641 ACATGTGGAACTTCTAGGCGCATCCAATACCTGCGAGGACTATCAACACTGCCAGG 5700
 5641 ACATGTGGAACTTCTAGGCGCATCCAATACCTGCGAGGACTATCAACACTGCCAGG 5700
 5701 GAACTGTGAGTGTTCATGAGGCTTCAATGCGCCCGCTCAGTGGCGCTCAGTGGCGCTCAGC 5760
 5701 GAACTGTGAGTGTTCATGAGGCTTCAATGCGCCCGCTCAGTGGCGCTCAGTGGCGCTCAGC 5760

1381	Qy	CATTAGCGGGCTCATTTGGGGCTCATGTTTGGCTTGGCCCTACTTCTCTATGCGAGGAGC	1440
1381	Db	CATTAGCGGGCTCATTTGGGGCTCATGTTTGGCTTGGCCCTACTTCTCTATGCGAGGAGC	1440
1441	Qy	GTGGGCAAAAGTCGTGTCATCTCTGTTGGCGCGCGGGGTGACGCGCGACCCATAC	1500
1441	Db	GTGGGCAAAAGTCGTGTCATCTCTGTTGGCGCGCGGGGTGACGCGCGACCCATAC	1500
1501	Qy	TGTTGGGGGTTCTGCCCGCGAGACCACCGGGCGCCTCACAGCTTATTTGACATGGGCC	1560
1501	Db	TGTTGGGGGTTCTGCCCGCGAGACCACCGGGCGCCTCACAGCTTATTTGACATGGGCC	1560
1561	Qy	CAGCGAAAAATCAGCTCGTTTAAACACCAATGGCAGCTGGCAGATCAACCGCACGGCCT	1620
1561	Db	CAGCGAAAAATCAGCTCGTTTAAACACCAATGGCAGCTGGCAGATCAACCGCACGGCCT	1620
1621	Qy	GAACTGCAATGACTCCCTTGACACCGGCTTATTCGGCTCTCTGTTCTACACCAAGCTT	1680
1621	Db	GAACTGCAATGACTCCCTTGACACCGGCTTATTCGGCTCTCTGTTCTACACCAAGCTT	1680
1681	Qy	CAACTCGTCAGAGTGTCCCGAAGCATGTCCGCTCGCGCAGTATCGAGGCTTCCGGGT	1740
1681	Db	CAACTCGTCAGAGTGTCCCGAAGCATGTCCGCTCGCGCAGTATCGAGGCTTCCGGGT	1740
1741	Qy	GGGATGGGGGCCCTTGCAATATGAGGATAATGTCAACAATCCAGAGATATGAGACCCCTA	1800
1741	Db	GGGATGGGGGCCCTTGCAATATGAGGATAATGTCAACAATCCAGAGATATGAGACCCCTA	1800
1801	Qy	TTGCTGGCACTACCCACCAAGGCAGTGTGGCGTGTCTCGCGAAGACTGTGTGGGCC	1860
1801	Db	TTGCTGGCACTACCCACCAAGGCAGTGTGGCGTGTCTCGCGAAGACTGTGTGGGCC	1860
1861	Qy	AGTGTACTGTTTACCCCGCAGCCAGTGTAGTGGGCAAGCCGACAGGCTTGGAGCGCC	1920
1861	Db	AGTGTACTGTTTACCCCGCAGCCAGTGTAGTGGGCAAGCCGACAGGCTTGGAGCGCC	1920
1921	Qy	CACTTACAGTGGGGGAGAAATGACAGATGTCTTCTATTGAAACAGCACTCGACCAAC	1980
1921	Db	CACTTACAGTGGGGGAGAAATGACAGATGTCTTCTATTGAAACAGCACTCGACCAAC	1980
1981	Qy	GCTGGGTCATGGTTTGGCTGACGTGGATGAATCTTCTGGGCTACACCAAGACTTGGG	2040
1981	Db	GCTGGGTCATGGTTTGGCTGACGTGGATGAATCTTCTGGGCTACACCAAGACTTGGG	2040
2041	Qy	CGCACCACTCGCTACTAGAGCTGACTTCAAGCGCAGCAGGCACTGTTGTGGCCCAAC	2100
2041	Db	CGCACCACTCGCTACTAGAGCTGACTTCAAGCGCAGCAGGCACTGTTGTGGCCCAAC	2100
2101	Qy	GGACTGTTTTAGGAAGCATCTGTATACCACTTACCTCAAAATGCGGCTCTGGGCCCTGGCT	2160
2101	Db	GGACTGTTTTAGGAAGCATCTGTATACCACTTACCTCAAAATGCGGCTCTGGGCCCTGGCT	2160
2161	Qy	CACGCAAGTGCTGATCGACTACCCCTACAGGCTCTGGCATTACCCCTGACAGTTAA	2220
2161	Db	CACGCAAGTGCTGATCGACTACCCCTACAGGCTCTGGCATTACCCCTGACAGTTAA	2220
2221	Qy	CTATACCATCTTCAAAATAGGAATGATGTGGAGGGGTTGACACAGGCTCACGGCTGC	2280
2221	Db	CTATACCATCTTCAAAATAGGAATGATGTGGAGGGGTTGACACAGGCTCACGGCTGC	2280
2281	Qy	ATGCNAATTTCACTCGTGGGATCGTTGCAACTTGGAGGACAGAGACAGAACTCACTGTC	2340
2281	Db	ATGCNAATTTCACTCGTGGGATCGTTGCAACTTGGAGGACAGAGACAGAACTCACTGTC	2340
2341	Qy	TCCTTTGTGTGCACTCCACCAAGGAATGGGCAATTTTACCTTGTCTTATCTCGGACCTGCC	2400
2341	Db	TCCTTTGTGTGCACTCCACCAAGGAATGGGCAATTTTACCTTGTCTTATCTCGGACCTGCC	2400
2401	Qy	CGCCTGTGCACTGGTCTTCTCCACCTCCACCAAAACATCGTGGAGCTCAAAATTCATGTA	2460
2401	Db	CGCCTGTGCACTGGTCTTCTCCACCTCCACCAAAACATCGTGGAGCTCAAAATTCATGTA	2460
2461	Qy	TGGCCTATCACCTGGCCCTCACAATAATCATCGTCCGATGGGAGTGGGTAATACTCTTATT	2520
2461	Db	TGGCCTATCACCTGGCCCTCACAATAATCATCGTCCGATGGGAGTGGGTAATACTCTTATT	2520
2521	Qy	CCTGCTCTTAGCGGAGCGCCAGGGTTTGGCCCTGCTTATGGAATGCTCATCTTGTGGGCA	2580
2521	Db	CCTGCTCTTAGCGGAGCGCCAGGGTTTGGCCCTGCTTATGGAATGCTCATCTTGTGGGCA	2580
2581	Qy	GGCGAAGCAGACCTATAGAGAGCTGTGTCTATGTCACGCTGGAGCGAGCTAGCTGCAA	2640
2581	Db	GGCGAAGCAGACCTATAGAGAGCTGTGTCTATGTCACGCTGGAGCGAGCTAGCTGCAA	2640
2641	Qy	TGGCTTCTTATATTTTGTCTATCTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2700
2641	Db	TGGCTTCTTATATTTTGTCTATCTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2700
2701	Qy	CCCTTTAGCTTACCTATTCCTCTCACTGGCCCTGTGGCTTCTTTAGCTTACTGCTCTAGCAAT	2760
2701	Db	CCCTTTAGCTTACCTATTCCTCTCACTGGCCCTGTGGCTTCTTTAGCTTACTGCTCTAGCAAT	2760
2761	Qy	GGCCCAACAGGCTTATGCTTATGACGATCTGTGATGCGCAGATAGGAGCGGCTCTGCT	2820
2761	Db	GGCCCAACAGGCTTATGCTTATGACGATCTGTGATGCGCAGATAGGAGCGGCTCTGCT	2820
2821	Qy	GCTGCTGCTGTGCTATCTTCTGACCTCGGGGAGCTATGCTTCAAGAGCTTCTCAGCGGCTTTT	2880
2821	Db	GCTGCTGCTGTGCTATCTTCTGACCTCGGGGAGCTATGCTTCAAGAGCTTCTCAGCGGCTTTT	2880
2881	Qy	GTGGTGGTGTGCTATCTTCTGACCTCGGGGAGCTATGCTTCAAGAGCTTCTCAGCGGCTTTT	2940
2881	Db	GTGGTGGTGTGCTATCTTCTGACCTCGGGGAGCTATGCTTCAAGAGCTTCTCAGCGGCTTTT	2940
2941	Qy	TATGCAAGTGTGCGGTGGCGCTGATGCGCATCATATGCGGCGCTTCTACCCAGG	3000
2941	Db	TATGCAAGTGTGCGGTGGCGCTGATGCGCATCATATGCGGCGCTTCTACCCAGG	3000
3001	Qy	TGTGGTGTGCTATTAACCAAGTGGCTCTTGGGGTGTGCTTGGGCTTCTTACCTCTTAAA	3060
3001	Db	TGTGGTGTGCTATTAACCAAGTGGCTCTTGGGGTGTGCTTGGGCTTCTTACCTCTTAAA	3060
3061	Qy	AGGTGCTTTGACCGCGTGTGCTTCTGAGGGCTCACGCTCTACTGAGGATGTGCAC	3120
3061	Db	AGGTGCTTTGACCGCGTGTGCTTCTGAGGGCTCACGCTCTACTGAGGATGTGCAC	3120
3121	Qy	CATGGCAAGCATCTCGCGGGGAGGTAAGTACGTCAGATGCGGCTACTAGCCCTTGGCAG	3180
3121	Db	CATGGCAAGCATCTCGCGGGGAGGTAAGTACGTCAGATGCGGCTACTAGCCCTTGGCAG	3180
3181	Qy	GTGACCTGGCACTTACATCTATGACCACTTATGTCGAGTGGGCTGCTAGTGG	3240
3181	Db	GTGACCTGGCACTTACATCTATGACCACTTATGTCGAGTGGGCTGCTAGTGG	3240
3241	Qy	CGTGGGAGCTGGCGGTGCGGCTTACCTTACCTTACCTTACCTTACCTT	3300
3241	Db	CGTGGGAGCTGGCGGTGCGGCTTACCTTACCTTACCTTACCTTACCTT	3300
3301	Qy	CATTGTCTGGGAGCGGAGACAGCTCTCTTGTGGGACATTTTACACGGACTTCCCGTGTGC	3360
3301	Db	CATTGTCTGGGAGCGGAGACAGCTCTCTTGTGGGACATTTTACACGGACTTCCCGTGTGC	3360
3361	Qy	CGCCGAGCTTGGTCTGGGAGGTTCTCTTGGCCAGCTGATGGCTATACCTCCAGGGGTG	3420
3361	Db	CGCCGAGCTTGGTCTGGGAGGTTCTCTTGGCCAGCTGATGGCTATACCTCCAGGGGTG	3420
3421	Qy	GAGTCTTCTCGCCCCCATCATCTCTTACGCCAGACAGACAGCTGGGCTTTTGGGACCAT	3480
3421	Db	GAGTCTTCTCGCCCCCATCATCTCTTACGCCAGACAGACAGCTGGGCTTTTGGGACCAT	3480
3481	Qy	AGTGGTGAGCATGACGGGGCGCGAAGAGCAGAAAGGCTGGGGAATTCAGGTCTGTGC	3540
3481	Db	AGTGGTGAGCATGACGGGGCGCGAAGAGCAGAAAGGCTGGGGAATTCAGGTCTGTGC	3540
3541	Qy	CACAGTCACTGAGTCTTCTCTCGGAAACATCTCTCGGGGTTTTGTGCACTGTCTACCA	3600
3541	Db	CACAGTCACTGAGTCTTCTCTCGGAAACATCTCTCGGGGTTTTGTGCACTGTCTACCA	3600

QY 9361 GCCTGCCGACCCCGCTTATGCTCTTTGGCCCTACTCCTACTTTTGTAGGGGTAGGCT 9420
 DB 9361 GCCTGCCGACCCCGCTTATGCTCTTTGGCCCTACTCCTACTTTTGTAGGGGTAGGCT 9420
 QY 9421 TTTCTTACTCCCGCTCGGTAGAGCGGACACATAGTACACTCCATAGTAATGTCC 9480
 DB 9421 TTTCTTACTCCCGCTCGGTAGAGCGGACACATAGTACACTCCATAGTAATGTCC 9480
 QY 9481 CTTTCTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9540
 DB 9481 CTTTCTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9540
 QY 9541 TTTTCTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9600
 DB 9541 TTTTCTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9600
 QY 9601 TACTTTCTTTCTTGTGGTCCATCTAGCCCTAGTACAGGCTAGTGTGAAAGTCCGT 9660
 DB 9601 TACTTTCTTTCTTGTGGTCCATCTAGCCCTAGTACAGGCTAGTGTGAAAGTCCGT 9660
 QY 9661 GAGCCGATGACTGCAGAGAGTCCGTAAGTCTCTCTGAGATCATGT 9711
 DB 9661 GAGCCGATGACTGCAGAGAGTCCGTAAGTCTCTCTGAGATCATGT 9711

RESULT 3
 AX057395
 LOCUS AX057395 9711 bp DNA linear PAT 17-JAN-2001
 DEFINITION Sequence 4 from Patent WO0075337.
 ACCESSION AX057395
 VERSION AX057395.1 GI:12310134
 KEYWORDS
 SOURCE Hepatitis C virus
 ORGANISM Hepatitis C virus
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.

REFERENCE 1
 AUTHORS Bukh, J., Yanagi, M., Emerson, S.U. and Purcell, R.H.
 TITLE Infectious cdna clone of gb virus b and uses thereof
 JOURNAL Patent: WO 0075337-A 4 14-DEC-2000;
 THE GOVERNMENT OF THE UNITED STATES OF AMERICA (US)
 FEATURES
 Location/Qualifiers
 1..9711
 /organism="Hepatitis C virus"
 /mol_type="genomic DNA"
 /db_xref="taxon:111103"
 BASE COUNT 1987 a 2853 c 2668 g 2203 t
 ORIGIN

Query Match 100.0%; Score 9711; DB 6; Length 9711;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 9711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCCGCCCTTATAGGGGACACTCCGCATGAATCACTCCCTCTGTGAGGAATCACTGT 60
 DB 1 ACCCGCCCTTATAGGGGACACTCCGCATGAATCACTCCCTCTGTGAGGAATCACTGT 60
 QY 61 CTTTCAGCAAGAGCGTCTAGCCATGGCTTGTATGATGTGCTACAGCCTCCAGGCC 120
 DB 61 CTTTCAGCAAGAGCGTCTAGCCATGGCTTGTATGATGTGCTACAGCCTCCAGGCC 120
 QY 121 CCCCTCCCGGAGAGCCATAGTGGTCTCGGAACCGGTAGTACACCGGAATTCGCGG 180
 DB 121 CCCCTCCCGGAGAGCCATAGTGGTCTCGGAACCGGTAGTACACCGGAATTCGCGG 180
 QY 181 AAGACTGGGTCTTTCTTGGATAAACCCACTCTATGCCCGGCATTTGGGCGTGCCTCCG 240
 DB 181 AAGACTGGGTCTTTCTTGGATAAACCCACTCTATGCCCGGCATTTGGGCGTGCCTCCG 240
 QY 241 CAAGACTGTAGCCGAGTACCGTGGTTCGGAAGCGCTTGTGGTACTGTGCTGATAGG 300
 DB 241 CAAGACTGTAGCCGAGTACCGTGGTTCGGAAGCGCTTGTGGTACTGTGCTGATAGG 300

QY 301 TGCTTTGCAGTGCCTCCGAGGTCTCTAGACCGGTGACCATGAGCAAAATCTTAAACC 360
 DB 301 TGCTTTGCAGTGCCTCCGAGGTCTCTAGACCGGTGACCATGAGCAAAATCTTAAACC 360
 QY 361 TCAAAAGAAAACAAAAGAAAACCAACCGTCCGCCCAACAAGAGTTAAGTTTCCGGCGG 420
 DB 361 TCAAAAGAAAACAAAAGAAAACCAACCGTCCGCCCAACAAGAGTTAAGTTTCCGGCGG 420
 QY 421 CGGCAGAGTCTGTGGCGAGTATACCTTGTGCGCGGAGGGCCAGGTGGGTGTGG 480
 DB 421 CGGCAGAGTCTGTGGCGAGTATACCTTGTGCGCGGAGGGCCAGGTGGGTGTGG 480
 QY 481 CGCGACAAGAGAGCTTCGAGCGGTCCAGCCAGTGTGAGGGCGCAGCCCATCCCTAA 540
 DB 481 CGCGACAAGAGAGCTTCGAGCGGTCCAGCCAGTGTGAGGGCGCAGCCCATCCCTAA 540
 QY 541 AGATCGGCGCTCCACTGTCGCAAAATCTGGGGAACACAGGATACCCCTGCCCCATACGG 600
 DB 541 AGATCGGCGCTCCACTGTCGCAAAATCTGGGGAACACAGGATACCCCTGCCCCATACGG 600
 QY 601 GAAATGAGGAGTCTGGCTCGGCGAGGATGCTCTGTCCCCCGAGGTTCCTGCTCTTTG 660
 DB 601 GAAATGAGGAGTCTGGCTCGGCGAGGATGCTCTGTCCCCCGAGGTTCCTGCTCTTTG 660
 QY 661 GGGCCCCAATGACCCCGGCATAGGTTCGCAAACTGTGGGTAAAGTCAATGATACCTAAC 720
 DB 661 GGGCCCCAATGACCCCGGCATAGGTTCGCAAACTGTGGGTAAAGTCAATGATACCTAAC 720
 QY 721 GTGCGGCTTTGCGACCTCATGGGTATACCTCTGTGCGCGCCCGCTCGGGCGGT 780
 DB 721 GTGCGGCTTTGCGACCTCATGGGTATACCTCTGTGCGCGCCCGCTCGGGCGGT 780
 QY 781 CGCAGAGCTCTCGCGCATGGGTGAGAGTCTTGAGAGCGGGGTAAATTTTGAACAGG 840
 DB 781 CGCAGAGCTCTCGCGCATGGGTGAGAGTCTTGAGAGCGGGGTAAATTTTGAACAGG 840
 QY 841 GAACTTACCCGGTGTCTCTTTCTATCTTCTGTGCGCCCTGCTGCTGCATCACCAC 900
 DB 841 GAACTTACCCGGTGTCTCTTTCTATCTTCTGTGCGCCCTGCTGCTGCATCACCAC 900
 QY 901 CCGGTCTCCGCTGCGAAGTGAAGAACATCAGTACCGGCTACATGTGAGTAAACGACTG 960
 DB 901 CCGGTCTCCGCTGCGAAGTGAAGAACATCAGTACCGGCTACATGTGAGTAAACGACTG 960
 QY 961 CACCAATGACAGCATTTACCTGGCAGCTCCAGGCTGCTCTCTCCAGTCCCGGGTGGT 1020
 DB 961 CACCAATGACAGCATTTACCTGGCAGCTCCAGGCTGCTCTCTCCAGTCCCGGGTGGT 1020
 QY 1021 CCGGTGCGAGAAAGTGGGGAATGCAATCTCAGTGTGGATACCGGTCTCACCGAATGTGC 1080
 DB 1021 CCGGTGCGAGAAAGTGGGGAATGCAATCTCAGTGTGGATACCGGTCTCACCGAATGTGC 1080
 QY 1081 CGTCAGCGGCGCGCCCTCAGCAGGGGTTCGCGACGACATCGACATGTTGTGAT 1140
 DB 1081 CGTCAGCGGCGCGCCCTCAGCAGGGGTTCGCGACGACATCGACATGTTGTGAT 1140
 QY 1141 GTCGCGCAGCTGTCTGCTCTGCTTCTAGTGGGGAACCTCTGCGGTGGGTGATGCTCG 1200
 DB 1141 GTCGCGCAGCTGTCTGCTCTGCTTCTAGTGGGGAACCTCTGCGGTGGGTGATGCTCG 1200
 QY 1201 AGCCCAATGTTTCTTCTGCTGCGCGACACACCTGTTTGTCCAAAGTTCGAATGTCTC 1260
 DB 1201 AGCCCAATGTTTCTTCTGCTGCGCGACACACCTGTTTGTCCAAAGTTCGAATGTCTC 1260
 QY 1261 CATCTACCTGTTACCATCTGGAACACCGCATGCGCATGGGACATGATGAACTGCTC 1320
 DB 1261 CATCTACCTGTTACCATCTGGAACACCGCATGCGCATGGGACATGATGAACTGCTC 1320
 QY 1321 GCCCAGGCTACCATGATCTTGGCGTACGCAATGCGGTGCCCGAGGTCAATTATAGAT 1380
 DB 1321 GCCCAGGCTACCATGATCTTGGCGTACGCAATGCGGTGCCCGAGGTCAATTATAGAT 1380

QY	4981	CAAACGCGCGTTTCCCTGTGTGCAAGACCACTTGTAGTTTGGGAGGCGAGTTTTCAC	5040	QY	6061	CGGACACGTCGGGACCGGGGAGCGCGTCCAAATGGATGAATAGACTCATTTGCTTTTGC	6120
DB	4981	CAAACGCGCGTTTCCCTGTGTGCAAGACCACTTGTAGTTTGGGAGGCGAGTTTTCAC	5040	DB	6061	CGGACACGTCGGGACCGGGGAGCGCGTCCAAATGGATGAATAGACTCATTTGCTTTTGC	6120
QY	5041	CGGCGCTCACACATAGATGCCACTTCTTTTCCCAAAACAAGCAATCGGGGAAATTTT	5100	QY	6121	TTCCAGAGGAATCACGTCGCGCCCAACCACTAAGTGAAGGAGTGGATGGTCCGACGG	6180
DB	5041	CGGCGCTCACACATAGATGCCACTTCTTTTCCCAAAACAAGCAATCGGGGAAATTTT	5100	DB	6121	TTCCAGAGGAATCACGTCGCGCCCAACCACTAAGTGAAGGAGTGGATGGTCCGACGG	6180
QY	5101	CGCATATTAACAGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCT	5160	QY	6181	TGTGACCCCAACTTACTTGTGCTCTTACCATAAACAGGCTGTCTCAGAAAGACTCCCAACTG	6240
DB	5101	CGCATATTAACAGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCTACAGGCT	5160	DB	6181	TGTGACCCCAACTTACTTGTGCTCTTACCATAAACAGGCTGTCTCAGAAAGACTCCCAACTG	6240
QY	5161	CTGGGACGTCATGTGAAGTGTGTGACTCGACTCAAGCCCACTCGTGGGCGCCACACC	5220	QY	6241	GATTAATGAGGACTGCCCCCATCCCATGCGCGGCTCGTGGCTCGCGGATGTGTGGACTG	6300
DB	5161	CTGGGACGTCATGTGAAGTGTGTGACTCGACTCAAGCCCACTCGTGGGCGCCACACC	5220	DB	6241	GATTAATGAGGACTGCCCCCATCCCATGCGCGGCTCGTGGCTCGCGGATGTGTGGACTG	6300
QY	5221	TCTCTGTACCGTCTGGGCTCTGTATACCAAGAGGTCACCCCTCACACATCCCGTGACGAA	5280	QY	6301	GTTTGGACCACTTCTTAAAGATTTTAAATTTGGCTGACCTCCAAATTTATTTCCCAAGAT	6360
DB	5221	TCTCTGTACCGTCTGGGCTCTGTATACCAAGAGGTCACCCCTCACACATCCCGTGACGAA	5280	DB	6301	GTTTGGACCACTTCTTAAAGATTTTAAATTTGGCTGACCTCCAAATTTATTTCCCAAGAT	6360
QY	5281	ATACATGCGCACCTGATGCAAGCCGACCTTGAGGTCATGACGACGACATGGGCTCTTGGC	5340	QY	6361	GCCGCGCTCCCTTGTCTCTGTCAAAAGGGGTACAAAGGGGTGTGGGCGGCGACTGG	6420
DB	5281	ATACATGCGCACCTGATGCAAGCCGACCTTGAGGTCATGACGACGACATGGGCTCTTGGC	5340	DB	6361	GCCGCGCTCCCTTGTCTCTGTCAAAAGGGGTACAAAGGGGTGTGGGCGGCGACTGG	6420
QY	5341	AGGGGAGTCTTGGCGCGCTGCGCGCTATTCCTTGGCGACCGGGTGTGTGATCAT	5400	QY	6421	CATCATGACACACGCTGTCTTGGCGCGCAATATCTCTGGCAATGTCCGCTTGGGCTC	6480
DB	5341	AGGGGAGTCTTGGCGCGCTGCGCGCTATTCCTTGGCGACCGGGTGTGTGATCAT	5400	DB	6421	CATCATGACACACGCTGTCTTGGCGCGCAATATCTCTGGCAATGTCCGCTTGGGCTC	6480
QY	5401	CGGCGCTTGCATTAACAGGAGCGCTGTGCGCGGACCAAGGAGTCTCTCTATGA	5460	QY	6481	CATGAGAAATCAAGGGGCTTAAAGCTGATGAATATCTGGCAGGGAACCTTTCTTATCAA	6540
DB	5401	CGGCGCTTGCATTAACAGGAGCGCTGTGCGCGGACCAAGGAGTCTCTCTATGA	5460	DB	6481	CATGAGAAATCAAGGGGCTTAAAGCTGATGAATATCTGGCAGGGAACCTTTCTTATCAA	6540
QY	5461	GGCTTTGTATGATGAGGAGATGCTCTTGGCGCGCTTCAATTTGAAGAGGGGCGAGG	5520	QY	6541	TTGTATACGAGGAGGCGAGTGGTGGCGAAACCGCGCGCAAACTTTAAGGTGCGCATCTG	6600
DB	5461	GGCTTTGTATGATGAGGAGATGCTCTTGGCGCGCTTCAATTTGAAGAGGGGCGAGG	5520	DB	6541	TTGTATACGAGGAGGCGAGTGGTGGCGAAACCGCGCGCAAACTTTAAGGTGCGCATCTG	6600
QY	5521	GATAGCGGAGTCTCAAGTCAAGATCAAGGCTTATTGACGAGGCTTCCAAACAAGC	5580	QY	6601	GAGGTGCGCGCTCAGAGTACGCGGAGTGAACGAGGTCATACCACTACATAAC	6660
DB	5521	GATAGCGGAGTCTCAAGTCAAGATCAAGGCTTATTGACGAGGCTTCCAAACAAGC	5580	DB	6601	GAGGTGCGCGCTCAGAGTACGCGGAGTGAACGAGGTCATACCACTACATAAC	6660
QY	5581	TCAGACATACCAACCACTGTGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGG	5640	QY	6661	AGGACTACCACTGATTAACCTTGAAGTCTTCCGCACTACCTTCTTCTTCTTCTTCTTCT	6720
DB	5581	TCAGACATACCAACCACTGTGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGG	5640	DB	6661	AGGACTACCACTGATTAACCTTGAAGTCTTCCGCACTACCTTCTTCTTCTTCTTCTTCT	6720
QY	5641	ACACATGTCGAACTTCAATAGCGGATCAATACCTCGAGGACTATCAACACTGCCAGG	5700	QY	6721	CTGGTGAACGAGTGCAGATCCATAGTTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	6780
DB	5641	ACACATGTCGAACTTCAATAGCGGATCAATACCTCGAGGACTATCAACACTGCCAGG	5700	DB	6721	CTGGTGAACGAGTGCAGATCCATAGTTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	6780
QY	5701	GAACCTGCACTAGTGTTCATGATGCGTTCAGTGGCGGCTTCAACGAGTCCGCTGTCAAC	5760	QY	6781	TGAGTCTCTGCTTCTGCGTGGGCTTAAATTTGTCTGCGGGTCCCAAGCTTCTTCTGCGA	6840
DB	5701	GAACCTGCACTAGTGTTCATGATGCGTTCAGTGGCGGCTTCAACGAGTCCGCTGTCAAC	5760	DB	6781	TGAGTCTCTGCTTCTGCGTGGGCTTAAATTTGTCTGCGGGTCCCAAGCTTCTTCTGCGA	6840
QY	5761	AAGCACCACTATCTTCTCAACATTTTGGGGGCTGGCTAGCATCCCAATTTGCACACC	5820	QY	6841	CCCTGAACCCGACACAGAGTATGATGTCATGTGTAAACAGATCCATCTCATATCACGCG	6900
DB	5761	AAGCACCACTATCTTCTCAACATTTTGGGGGCTGGCTAGCATCCCAATTTGCACACC	5820	DB	6841	CCCTGAACCCGACACAGAGTATGATGTCATGTGTAAACAGATCCATCTCATATCACGCG	6900
QY	5821	CGGCGGGGCGACTGGGCTTGTGTAGTGGCTTGTGAGGAGTGGGAGTGGGAGTATAGG	5880	QY	6901	GGAGACTGCGCGGCGGCTTGTAGCGGGGCTTGTAGCGGGGCTTGTAGCGGGGCTTGTAG	6960
DB	5821	CGGCGGGGCGACTGGGCTTGTGTAGTGGCTTGTGAGGAGTGGGAGTGGGAGTATAGG	5880	DB	6901	GGAGACTGCGCGGCGGCTTGTAGCGGGGCTTGTAGCGGGGCTTGTAGCGGGGCTTGTAG	6960
QY	5881	CTTAGGTAAAGTGTAGTGAATCTTGGCAGGATGTGTGGGCTTGTGGGCTTGTGGGCT	5940	QY	6961	GAGCAGCTATCGGACCACTCGCTGCGAGCACAAGTGTGACCAACCAAGGCTTATGA	7020
DB	5881	CTTAGGTAAAGTGTAGTGAATCTTGGCAGGATGTGTGGGCTTGTGGGCTTGTGGGCT	5940	DB	6961	GAGCAGCTATCGGACCACTCGCTGCGAGCACAAGTGTGACCAACCAAGGCTTATGA	7020
QY	5941	CGTGCATTAACAGATCATGTCTGGGAGAGCCCTTCAATGGAGGATGTCTCAACTTGTCT	6000	QY	7021	TGTGACATGTGTGATGCTAACTTGTCTGAGGAGGCTGATGTGACTCGGATAGAGTCTGG	7080
DB	5941	CGTGCATTAACAGATCATGTCTGGGAGAGCCCTTCAATGGAGGATGTCTCAACTTGTCT	6000	DB	7021	TGTGACATGTGTGATGCTAACTTGTCTGAGGAGGCTGATGTGACTCGGATAGAGTCTGG	7080
QY	6001	GCTTGAATTTCTCTTCTGGGCTTGTGTAGTGGGAGTCACTTGGGCGGCGCTTCTGGG	6060	QY	7081	GTCGAAAGTGTCTTCTGAGGCTTCTGAGCCCAATGTGTGAGGAGGAGGAGGAGGAGGAG	7140
DB	6001	GCTTGAATTTCTCTTCTGGGCTTGTGTAGTGGGAGTCACTTGGGCGGCGCTTCTGGG	6060	DB	7081	GTCGAAAGTGTCTTCTGAGGCTTCTGAGCCCAATGTGTGAGGAGGAGGAGGAGGAGGAG	7140
QY	6061	GCTTGAATTTCTCTTCTGGGCTTGTGTAGTGGGAGTCACTTGGGCGGCGCTTCTGGG	6060	QY	7141	GCCTTCGATACCATCAGAAATATGCTCTCCCAAGAGAGGTTCCCAAGGCTTTTACCGGC	7200

Db	2761	GCCCCAAGGCTTATGCTTATGAGCATCTGTGTCATGGCCAGATAGGACGGCTCTGCT	2820	Db	3841	GTCTCTCAGAGGCGCGGTGTATGCCCCAGGGGCCACGCTGTGTGGAGTCTTCCGGGCGAGC	3900
Qy	2821	GGTAATGATCAGCTCTTTACTCTCACCCCGGGTATAGACCTTCTCAGCGGTTTTT	2880	Qy	3901	TGTGTGCTCTCGGGCGGTGGCTAACTCCATAGATTTTCACTCCCGTTGAGACACTCGACAT	3960
Db	2821	GGTAATGATCAGCTCTTTACTCTCACCCCGGGTATAGACCTTCTCAGCGGTTTTT	2880	Db	3901	TGTGTGCTCTCGGGCGGTGGCTAACTCCATAGATTTTCACTCCCGTTGAGACACTCGACAT	3960
Qy	2881	GTGGTGTGTGTCTATCTTCTGACCCCTGGGGGAAGCTATGGTCCAGAGTGGGCAACACC	2940	Qy	3961	CGTACAGCGGTCCCCACCTTTAGTGAACAAGCAACACCTGTGTGCCCCAGACCTTA	4020
Db	2881	GTGGTGTGTGTCTATCTTCTGACCCCTGGGGGAAGCTATGGTCCAGAGTGGGCAACACC	2940	Db	3961	CGTACAGCGGTCCCCACCTTTAGTGAACAAGCAACACCTGTGTGCCCCAGACCTTA	4020
Qy	2941	TATGCAAGTCCGCGGTGGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT	3000	Qy	4021	TCAGTCCGGTACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	4080
Db	2941	TATGCAAGTCCGCGGTGGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT	3000	Db	4021	TCAGTCCGGTACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	4080
Qy	3001	TGTGGTGTGTGATACAAAGTGGCTTCTTGGCGGTGCTTGGCGCTTACCTCTTAA	3060	Qy	4081	ATATGCTGTCTCAGGGGTATAAAGTGTCTAGTGTCTTAAATCCCTCAGTGGCTGCCACCTGG	4140
Db	3001	TGTGGTGTGTGATACAAAGTGGCTTCTTGGCGGTGCTTGGCGCTTACCTCTTAA	3060	Db	4081	ATATGCTGTCTCAGGGGTATAAAGTGTCTAGTGTCTTAAATCCCTCAGTGGCTGCCACCTGG	4140
Qy	3061	AGGTGCTTTGACCGCGTGCCTTACTTCTGTCAGGGCTCAGCTTACTGAGGATGTGCAC	3120	Qy	4141	GTTTGGGGGTACTTGTCTTAAGGCACATGGGCATCAATCCCAATTTAGGACTGGAGTCA	4200
Db	3061	AGGTGCTTTGACCGCGTGCCTTACTTCTGTCAGGGCTCAGCTTACTGAGGATGTGCAC	3120	Db	4141	GTTTGGGGGTACTTGTCTTAAGGCACATGGGCATCAATCCCAATTTAGGACTGGAGTCA	4200
Qy	3121	CATGGCAAGGCATCTCGCGGGGCGAGGTACGTCCAGATGGCGTACTAGCCCTTGGCAG	3180	Qy	4201	GACTGTGACGACCGGGGCGCCCATCAGCTACTCCACATATGGCAAAATTCCTCGCCGATGG	4260
Db	3121	CATGGCAAGGCATCTCGCGGGGCGAGGTACGTCCAGATGGCGTACTAGCCCTTGGCAG	3180	Db	4201	GACTGTGACGACCGGGGCGCCCATCAGCTACTCCACATATGGCAAAATTCCTCGCCGATGG	4260
Qy	3181	GTGGACTGGCACTTACATCTATGACCACTTACATCTTCACTTCACTTCACTTCACTTCACT	3240	Qy	4261	GGGCTGTGCGGGCGGCGCTTACGACATCATATGATGATGATGATGATGATGATGATGATGAT	4320
Db	3181	GTGGACTGGCACTTACATCTATGACCACTTACATCTTCACTTCACTTCACTTCACTTCACT	3240	Db	4261	GGGCTGTGCGGGCGGCGCTTACGACATCATATGATGATGATGATGATGATGATGATGATGAT	4320
Qy	3241	CCTGCGGGACCTGCGCGTTCGAGCTTATCATCTTCACTTCACTTCACTTCACTTCACTTCACT	3300	Qy	4321	TACCAACATCTTGGCATCGGAACAGTCTTGTGATCAAGCAGACAGCTGGGGTCA	4380
Db	3241	CCTGCGGGACCTGCGCGTTCGAGCTTATCATCTTCACTTCACTTCACTTCACTTCACTTCACT	3300	Db	4321	TACCAACATCTTGGCATCGGAACAGTCTTGTGATCAAGCAGACAGCTGGGGTCA	4380
Qy	3301	CATTGTCTGGGAGCGGAGACAGCTGTCTTGGGGAATTTTACAGGCACTTCCCGTCTC	3360	Qy	4381	AACTGTGTGCTTACAGCTACGCTACGCTACGCTACGCTACGCTACGCTACGCTACGCTACG	4440
Db	3301	CATTGTCTGGGAGCGGAGACAGCTGTCTTGGGGAATTTTACAGGCACTTCCCGTCTC	3360	Db	4381	AACTGTGTGCTTACAGCTACGCTACGCTACGCTACGCTACGCTACGCTACGCTACGCTACG	4440
Qy	3361	CGCCGAGCTTGGTCCGAGGTCTCTTCTGCGCCAGCTGATGGCTTATACCTCCAAAGGGGTG	3420	Qy	4441	GGAGTGTGCGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	4500
Db	3361	CGCCGAGCTTGGTCCGAGGTCTCTTCTGCGCCAGCTGATGGCTTATACCTCCAAAGGGGTG	3420	Db	4441	GGAGTGTGCGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	4500
Qy	3421	GAGTCTTCTCGCCCGCATCTGCTTACGCGGAGCAGCTGCGGCTTGGGCACTT	3480	Qy	4501	TTACATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	4560
Db	3421	GAGTCTTCTCGCCCGCATCTGCTTACGCGGAGCAGCTGCGGCTTGGGCACTT	3480	Db	4501	TTACATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	4560
Qy	3481	AGTGTGAGCATGACGGGGCGGCAAGACAGAAACAGGCTGGGGAATTCAGGTCCTGTC	3540	Qy	4561	CGCGGGCGGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	4620
Db	3481	AGTGTGAGCATGACGGGGCGGCAAGACAGAAACAGGCTGGGGAATTCAGGTCCTGTC	3540	Db	4561	CGCGGGCGGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	4620
Qy	3541	CACAGTCACTCAGTCTTCTCGGAAATCCATCTCGGGGGTCTTGGGACTGTCTACCA	3600	Qy	4621	CTCCGTAATACCAACTCAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	4680
Db	3541	CACAGTCACTCAGTCTTCTCGGAAATCCATCTCGGGGGTCTTGGGACTGTCTACCA	3600	Db	4621	CTCCGTAATACCAACTCAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	4680
Qy	3601	TGGAGCTGGCAACAGACTCTGGGCGGCTCAGCGGGTCCGCTCAGCAGATGTACTCCAG	3660	Qy	4681	GTATATCGGGGACTTTGACTCGGTGATCGACTGCAAGTAGGGGTCACTCAAGTGTGTAGA	4740
Db	3601	TGGAGCTGGCAACAGACTCTGGGCGGCTCAGCGGGTCCGCTCAGCAGATGTACTCCAG	3660	Db	4681	GTATATCGGGGACTTTGACTCGGTGATCGACTGCAAGTAGGGGTCACTCAAGTGTGTAGA	4740
Qy	3661	TGCTGAGGGGACTTAGTGGGTGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	3720	Qy	4741	CTTCAGTGTAGACCCCACTTACCAATACCAATACCAATACCAATACCAATACCAATACCAATACCA	4800
Db	3661	TGCTGAGGGGACTTAGTGGGTGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	3720	Db	4741	CTTCAGTGTAGACCCCACTTACCAATACCAATACCAATACCAATACCAATACCAATACCAATACCA	4800
Qy	3721	CACGTGTGGAGCGGTGACCTTGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3780	Qy	4801	ACGTAGCCAGCGCGGGGTGCGCAACGAGTGGGGAAGACTGGGCAATTTATAGGTATGTTTC	4860
Db	3721	CACGTGTGGAGCGGTGACCTTGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3780	Db	4801	ACGTAGCCAGCGCGGGGTGCGCAACGAGTGGGGAAGACTGGGCAATTTATAGGTATGTTTC	4860
Qy	3781	AAGACGGGGGACAAAACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	3840	Qy	4861	CATGTGTAGGAGCGCTCAGGAAATGTTTGAAGTGTAGTGTCTGTGTGTGTGTGTGTGTGTGTGTGT	4920
Db	3781	AAGACGGGGGACAAAACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	3840	Db	4861	CATGTGTAGGAGCGCTCAGGAAATGTTTGAAGTGTAGTGTCTGTGTGTGTGTGTGTGTGTGTGTGT	4920
Qy	3841	GTCTCTCAGGAGGCGGTGTCTATGCCAGGGGCGACGCTGTGCGAGTCTTCCGGGCGAGC	3900	Qy	4921	AGGGGCGGCAATGTATGAGCTCACCAATCGGAGACCAACGCTCAGGCTCAGGCGGTATTTT	4980
Db	3841	GTCTCTCAGGAGGCGGTGTCTATGCCAGGGGCGACGCTGTGCGAGTCTTCCGGGCGAGC	3900	Db	4921	AGGGGCGGCAATGTATGAGCTCACCAATCGGAGACCAACGCTCAGGCTCAGGCGGTATTTT	4980

QY 601 GAATGAGGACTCGGCTGGGAGGATGGCTCTGTCCCCCGAGGTTCCGCTCCCTCTTG 660
 Db 601 GAATGAGGACTCGGCTGGGAGGATGGCTCTGTCCCCCGAGGTTCCGCTCCCTCTTG 660
 QY 661 GGGCCCCAATGATCCCCCGGATAGGTGCGGCAACGTTGGTAAAGTCACTGATACCCCTAAC 720
 Db 661 GGGCCCCAATGATCCCCCGGATAGGTGCGGCAACGTTGGTAAAGTCACTGATACCCCTAAC 720
 QY 721 GTGGGCTTTGCGGACTCATGGGTACATCCCTGTCTGTGGGCGCCCGCTCGGGGGGT 780
 Db 721 GTGGGCTTTGCGGACTCATGGGTACATCCCTGTCTGTGGGCGCCCGCTCGGGGGGT 780
 QY 781 CGCCAGAGCTCTCGGCAATGGGTGAGAGTCTCTGTGGGCGCCCGCTCGGGGGGT 840
 Db 781 CGCCAGAGCTCTCGGCAATGGGTGAGAGTCTCTGTGGGCGCCCGCTCGGGGGGT 840
 QY 841 GAATTTACCCGGTGTCTCTTTTCTATCTTCTGTGTGGGCGCCCGCTCGGGGGGT 900
 Db 841 GAATTTACCCGGTGTCTCTTTTCTATCTTCTGTGTGGGCGCCCGCTCGGGGGGT 900
 QY 901 CCGGCTCTCGGCTCGGAGTGAAGAACATCAGTACCGGCTACATGCTAAAGACTG 960
 Db 901 CCGGCTCTCGGCTCGGAGTGAAGAACATCAGTACCGGCTACATGCTAAAGACTG 960
 QY 961 CACCAATGACAGCATTTACCTGGCAGCTCCAGGCTGTCTCTCCAGCTCCCGGGTGGGT 1020
 Db 961 CACCAATGACAGCATTTACCTGGCAGCTCCAGGCTGTCTCTCCAGCTCCCGGGTGGGT 1020
 QY 1021 CCGGCTCGGAGAAAGTGGGAAATGCAATCTCAGTGTGGATACCGGCTCTACCGAAATGGG 1080
 Db 1021 CCGGCTCGGAGAAAGTGGGAAATGCAATCTCAGTGTGGATACCGGCTCTACCGAAATGGG 1080
 QY 1081 CGTGACGCGCGCGGCGCTCAGCAGGCTTGGGAGCGACATCGACATGCTGTGTAT 1140
 Db 1081 CGTGACGCGCGCGGCGCTCAGCAGGCTTGGGAGCGACATCGACATGCTGTGTAT 1140
 QY 1141 GTCCGCGACGCTCTGCTGCGCTCTACGCTGGGAGACCTCTGCGGTGGGATGATGCTGCG 1200
 Db 1141 GTCCGCGACGCTCTGCTGCGCTCTACGCTGGGAGACCTCTGCGGTGGGATGATGCTGCG 1200
 QY 1201 AGCCCAATGTTCAATGTTCTGCGCGGAGCAACCACTGTTGTTGCCAAGACTGCAATGCTC 1260
 Db 1201 AGCCCAATGTTCAATGTTCTGCGCGGAGCAACCACTGTTGTTGCCAAGACTGCAATGCTC 1260
 QY 1261 CATCTACCTGTTACCATCACTGGAACCGATGGCATGGGATGATGATGATGATGATGAT 1320
 Db 1261 CATCTACCTGTTACCATCACTGGAACCGATGGCATGGGATGATGATGATGATGATGAT 1320
 QY 1321 GCGCACGGCTTACCATGATCTTGGCGTACGCGATGCGTGTCCCGAGGCTCATTTATAGAT 1380
 Db 1321 GCGCACGGCTTACCATGATCTTGGCGTACGCGATGCGTGTCCCGAGGCTCATTTATAGAT 1380
 QY 1381 CATTTAGCGGGCTCATTTGGGGCTCATGTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1440
 Db 1381 CATTTAGCGGGCTCATTTGGGGCTCATGTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1440
 QY 1441 GTGGGCGAAGTGTGTTGATCTCTTGTGGGCGCGGGTGGAGCGCGGCGGCGGCGGCGG 1500
 Db 1441 GTGGGCGAAGTGTGTTGATCTCTTGTGGGCGCGGGTGGAGCGCGGCGGCGGCGGCGG 1500
 QY 1501 TGTGGGGGTGTGGCGCGGAGACCAACCGGGCGGCTTACAGCTTATTTGATGATGGGCGG 1560
 Db 1501 TGTGGGGGTGTGGCGCGGAGACCAACCGGGCGGCTTACAGCTTATTTGATGATGGGCGG 1560
 QY 1561 CAGGCGAAGAAATCCAGCTGTTAAACCAATGGAGCTGGGACATCAACCGGCGGCGGCT 1620
 Db 1561 CAGGCGAAGAAATCCAGCTGTTAAACCAATGGAGCTGGGACATCAACCGGCGGCGGCT 1620
 QY 1621 GAACTGCAATGACTCTTTCGACACCGGCTTATCGGGCTCTGTCTTACACCGGCGGCT 1680
 Db 1621 GAACTGCAATGACTCTTTCGACACCGGCTTATCGGGCTCTGTCTTACACCGGCGGCT 1680

QY 1681 CAACTCGTCCAGATGTCCCGAAACGATGTCCGCTGCGCGAGTATCGAGGCTTCCGGGT 1740
 Db 1681 CAACTCGTCCAGATGTCCCGAAACGATGTCCGCTGCGCGAGTATCGAGGCTTCCGGGT 1740
 QY 1741 GGGATGGGGCGGCTTGCATATGAGGATAATGTCCCAATCCAGAGGATATGAGACCCCTA 1800
 Db 1741 GGGATGGGGCGGCTTGCATATGAGGATAATGTCCCAATCCAGAGGATATGAGACCCCTA 1800
 QY 1801 TTGCTGGCACTACCCACCAAGGAGTGTGGGCTGGTCTCCGGAAGACTGTGTGTGGGCC 1860
 Db 1801 TTGCTGGCACTACCCACCAAGGAGTGTGGGCTGGTCTCCGGAAGACTGTGTGTGGGCC 1860
 QY 1861 AGTGTACTGTTTACCCCGGAGTGTAGTGGGCAACCGGAGCTTGGAGCGCC 1920
 Db 1861 AGTGTACTGTTTACCCCGGAGTGTAGTGGGCAACCGGAGCTTGGAGCGCC 1920
 QY 1921 CACTTACACGTGGGGGAGAAATGAGACAGATGTCTTCTATTTGAACAGCACTCGACCAAC 1980
 Db 1921 CACTTACACGTGGGGGAGAAATGAGACAGATGTCTTCTATTTGAACAGCACTCGACCAAC 1980
 QY 1981 GCTGGGGTCAATGTTTCCGCTGCACTGAGCTGACTTCAACCGGAGCAACCGGAGCTTGGG 2040
 Db 1981 GCTGGGGTCAATGTTTCCGCTGCACTGAGCTGACTTCAACCGGAGCAACCGGAGCTTGGG 2040
 QY 2041 CGCACCACTTCCGCTTACTAGAGCTGACTTCAACCGGAGCAACCGGAGCTTGGGCTTGG 2100
 Db 2041 CGCACCACTTCCGCTTACTAGAGCTGACTTCAACCGGAGCAACCGGAGCTTGGGCTTGG 2100
 QY 2101 GGACTGTTTGGAGAGATCTCTGATACCACTTACTTCAATCTGAGCTTGGGCTTGGGCTTGG 2160
 Db 2101 GGACTGTTTGGAGAGATCTCTGATACCACTTACTTCAATCTGAGCTTGGGCTTGGGCTTGG 2160
 QY 2161 CAGCCAAAGTGTCTGATCGACTACCTTCAACCGGAGCAACCGGAGCTTGGGCTTGGG 2220
 Db 2161 CAGCCAAAGTGTCTGATCGACTACCTTCAACCGGAGCAACCGGAGCTTGGGCTTGGG 2220
 QY 2221 CTATACCATCTTCAAAATAGGATGTATGTGGGAGGTTGAGCAGAGCTTCAAGGCTTGG 2280
 Db 2221 CTATACCATCTTCAAAATAGGATGTATGTGGGAGGTTGAGCAGAGCTTCAAGGCTTGG 2280
 QY 2281 ATGCAATTTCACTCGTGGGATCGTTGCACTTGGAGGAGCAGAGCAAGTCACTGCT 2340
 Db 2281 ATGCAATTTCACTCGTGGGATCGTTGCACTTGGAGGAGCAGAGCAAGTCACTGCT 2340
 QY 2341 TCCCTTGTGGCACTCCACCGAATGGGCTTACCTTGTCTTACTCGGACCTTGGC 2400
 Db 2341 TCCCTTGTGGCACTCCACCGAATGGGCTTACCTTGTCTTACTCGGACCTTGGC 2400
 QY 2401 CGCCTTGTGCACTGCTTCTTCCACCTCCACCAAAACATCGTGGAGCTTCAATTCATGTA 2460
 Db 2401 CGCCTTGTGCACTGCTTCTTCCACCTCCACCAAAACATCGTGGAGCTTCAATTCATGTA 2460
 QY 2461 TGGCCTTACCTGCTTCAAAATACATCGTCCGATGGGAGTGGGTAATCTTATTT 2520
 Db 2461 TGGCCTTACCTGCTTCAAAATACATCGTCCGATGGGAGTGGGTAATCTTATTT 2520
 QY 2521 CCTGCTTACGCGAGCGGCTTGGGCTGCTTATGATGCTCATCTTGTGTGGGCGA 2580
 Db 2521 CCTGCTTACGCGAGCGGCTTGGGCTGCTTATGATGCTCATCTTGTGTGGGCGA 2580
 QY 2581 GGGCGAAGCAGCACTAGAGAGCTGCTCATCTTGCACGCTGCGAGCGGAGCTAGCTGCA 2640
 Db 2581 GGGCGAAGCAGCACTAGAGAGCTGCTCATCTTGCACGCTGCGAGCGGAGCTAGCTGCA 2640
 QY 2641 TGGCTTCTATTTTGTGCTTCTTGTGGGCTTGTGTAATCAAGGCTGGGTAGT 2700
 Db 2641 TGGCTTCTATTTTGTGCTTCTTGTGGGCTTGTGTAATCAAGGCTGGGTAGT 2700
 QY 2701 CCCCCTAGCTTACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2760
 Db 2701 CCCCCTAGCTTACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2760
 QY 2761 GCGGCGAAGGCTTATGCTTATGAGCATCTGTGCTATGGGCGAGATAGGAGCGGCTTCTGCT 2820

8521	Db	8521	CACCACTAGCATGGGGAAACACATACATGCTACGTGAAAGCCTTAGCGGCTGTAAAGC	8580
8581	Qy	8581	TGCAAGGATAAATCGCGCCCAACAATGCTGGTATGCGCGCATGACTTGGTTGTCTCATCTCAGA	8640
	Qy	8581	TGCAAGGATAAATCGCGCCCAACAATGCTGGTATGCGCGCATGACTTGGTTGTCTCATCTCAGA	8640
	Db	8581	TGCAAGGATAAATCGCGCCCAACAATGCTGGTATGCGCGCATGACTTGGTTGTCTCATCTCAGA	8640
8641	Qy	8641	AAGCAGGGGACCGAGGAGGACGAGCGGAAACCTGAGAGCCTTCACGGAGGCTATGACCCAG	8700
	Db	8641	AAGCAGGGGACCGAGGAGGACGAGCGGAAACCTGAGAGCCTTCACGGAGGCTATGACCCAG	8700
	Qy	8641	AAGCAGGGGACCGAGGAGGACGAGCGGAAACCTGAGAGCCTTCACGGAGGCTATGACCCAG	8700
8701	Qy	8701	GTATTCTGCCCTCTCTGGTGTACCCGCCACAGACGGAGTATGATCTGGAGCTGATAAACATC	8760
	Db	8701	GTATTCTGCCCTCTCTGGTGTACCCGCCACAGACGGAGTATGATCTGGAGCTGATAAACATC	8760
	Qy	8701	GTATTCTGCCCTCTCTGGTGTACCCGCCACAGACGGAGTATGATCTGGAGCTGATAAACATC	8760
8761	Qy	8761	TTTGCTCCTCAAAATGTGTCTGTGGCGCTGGGCCCCACAAGGCCCGCCGAGATACTACTGAC	8820
	Db	8761	TTTGCTCCTCAAAATGTGTCTGTGGCGCTGGGCCCCACAAGGCCCGCCGAGATACTACTGAC	8820
	Qy	8761	TTTGCTCCTCAAAATGTGTCTGTGGCGCTGGGCCCCACAAGGCCCGCCGAGATACTACTGAC	8820
8821	Qy	8821	CAGAGACCTTACCATCTCCAATCGCCCGGCTGCTGGGAAAACAGTTAGACACATCCCTGT	8880
	Db	8821	CAGAGACCTTACCATCTCCAATCGCCCGGCTGCTGGGAAAACAGTTAGACACATCCCTGT	8880
	Qy	8821	CAGAGACCTTACCATCTCCAATCGCCCGGCTGCTGGGAAAACAGTTAGACACATCCCTGT	8880
8881	Qy	8881	CAATTATCGCTGGGAAAACATCATCAGTACGCCCGCACCATATGGGCTCGCATGGTCCCT	8940
	Db	8881	CAATTATCGCTGGGAAAACATCATCAGTACGCCCGCACCATATGGGCTCGCATGGTCCCT	8940
	Qy	8881	CAATTATCGCTGGGAAAACATCATCAGTACGCCCGCACCATATGGGCTCGCATGGTCCCT	8940
8941	Qy	8941	GATGACACACTTCTTCTCCAATCTCATGGCTCAAGACACGCTGGACAGAACTCAACTT	9000
	Db	8941	GATGACACACTTCTTCTCCAATCTCATGGCTCAAGACACGCTGGACAGAACTCAACTT	9000
	Qy	8941	GATGACACACTTCTTCTCCAATCTCATGGCTCAAGACACGCTGGACAGAACTCAACTT	9000
9001	Qy	9001	TGAGATGTACGGAGCGGTGTACTCTGTGAGTCCCTTGGACCTCCAGCTTATTAATGAAAG	9060
	Db	9001	TGAGATGTACGGAGCGGTGTACTCTGTGAGTCCCTTGGACCTCCAGCTTATTAATGAAAG	9060
	Qy	9001	TGAGATGTACGGAGCGGTGTACTCTGTGAGTCCCTTGGACCTCCAGCTTATTAATGAAAG	9060
9061	Qy	9061	GTTACATGGGCTGACGCTTTTCTCTGCACACATACATCTCCACAGAACTGACAGGGT	9120
	Db	9061	GTTACATGGGCTGACGCTTTTCTCTGCACACATACATCTCCACAGAACTGACAGGGT	9120
	Qy	9061	GTTACATGGGCTGACGCTTTTCTCTGCACACATACATCTCCACAGAACTGACAGGGT	9120
9121	Qy	9121	GGCTTCAGCCCTCAGAAAACTTGGGGCGCCACCCCTCAGAGCGTGAAGAGCGGGCAGG	9180
	Db	9121	GGCTTCAGCCCTCAGAAAACTTGGGGCGCCACCCCTCAGAGCGTGAAGAGCGGGCAGG	9180
	Qy	9121	GGCTTCAGCCCTCAGAAAACTTGGGGCGCCACCCCTCAGAGCGTGAAGAGCGGGCAGG	9180
9181	Qy	9181	TGCAGTTCAGGGCGTCCCTCATCTCCGTGGGGGAGAGCGCGCTTGTGGCTCGATATCT	9240
	Db	9181	TGCAGTTCAGGGCGTCCCTCATCTCCGTGGGGGAGAGCGCGCTTGTGGCTCGATATCT	9240
	Qy	9181	TGCAGTTCAGGGCGTCCCTCATCTCCGTGGGGGAGAGCGCGCTTGTGGCTCGATATCT	9240
9241	Qy	9241	CTTCAATTGGGGCGTGAAGACCAAGCTCAAACTCACTCCTCAATGTCGGAGAGCGGCTCTCT	9300
	Db	9241	CTTCAATTGGGGCGTGAAGACCAAGCTCAAACTCACTCCTCAATGTCGGAGAGCGGCTCTCT	9300
	Qy	9241	CTTCAATTGGGGCGTGAAGACCAAGCTCAAACTCACTCCTCAATGTCGGAGAGCGGCTCTCT	9300
9301	Qy	9301	GGATTTATCCAGCTGGTTTCAACCGTCGGCGCGCGGGGCGACATTTATCACAGCTGTCT	9360
	Db	9301	GGATTTATCCAGCTGGTTTCAACCGTCGGCGCGCGGGGCGACATTTATCACAGCTGTCT	9360
	Qy	9301	GGATTTATCCAGCTGGTTTCAACCGTCGGCGCGCGGGGCGACATTTATCACAGCTGTCT	9360
9361	Qy	9361	GGGTGCCGACCCCGCTTATTTGCTCTTTGGCTTACTCTTCTTTTGTAGGGGTAGGCCCT	9420
	Db	9361	GGGTGCCGACCCCGCTTATTTGCTCTTTGGCTTACTCTTCTTTTGTAGGGGTAGGCCCT	9420
	Qy	9361	GGGTGCCGACCCCGCTTATTTGCTCTTTGGCTTACTCTTCTTTTGTAGGGGTAGGCCCT	9420
9421	Qy	9421	TTTCTACTCTCCCGCTCGGTAGAGGGGACACATTAGCTACACTCCATAGCTAACTGTCC	9480
	Db	9421	TTTCTACTCTCCCGCTCGGTAGAGGGGACACATTAGCTACACTCCATAGCTAACTGTCC	9480
	Qy	9421	TTTCTACTCTCCCGCTCGGTAGAGGGGACACATTAGCTACACTCCATAGCTAACTGTCC	9480
9481	Qy	9481	CTTT	9540
	Db	9481	CTTT	9540
	Qy	9481	CTTT	9540
9541	Qy	9541	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCTCTTTTCTCTTTTCTTTCTTACTTATTT	9600
	Db	9541	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCTCTTTTCTCTTTTCTTTCTTACTTATTT	9600
	Qy	9541	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCTCTTTTCTCTTTTCTTTCTTACTTATTT	9600
9601	Qy	9601	TACTTTCTTTCTGGTGGCTCCACTTTAGCCCTAGTCAAGGCTAGCTGTGAAAGTCCGT	9660
	Db	9601	TACTTTCTTTCTGGTGGCTCCACTTTAGCCCTAGTCAAGGCTAGCTGTGAAAGTCCGT	9660
	Qy	9601	TACTTTCTTTCTGGTGGCTCCACTTTAGCCCTAGTCAAGGCTAGCTGTGAAAGTCCGT	9660

6361 GCCGGCCTCCCTTTGCTCTCTGTCACAAAGGGGTACAAAGGGGTGTGGGGCCGGCACTGG 6420
 6361 GCCGGCCTCCCTTTGCTCTCTGTCACAAAGGGGTACAAAGGGGTGTGGGGCCGGCACTGG 6420
 6421 CATCATGACACACGGTGTCTTGGGGCGCAATATCTCTGGCAATGTCCGCTTGGGCTC 6480
 6421 CATCATGACACACGGTGTCTTGGGGCGCAATATCTCTGGCAATGTCCGCTTGGGCTC 6480
 6481 CATGAGAATCACGGGGCCCTAAGACCTGCAATGATATCTGCGAGGGAGCCCTTCCATCAA 6540
 6481 CATGAGAATCACGGGGCCCTAAGACCTGCAATGATATCTGCGAGGGAGCCCTTCCATCAA 6540
 6541 TTGTTACACGGAGGGCCAGTGTGCGGAAACCCCGGCCAAACTTTAAAGGTGCCATCTG 6600
 6541 TTGTTACACGGAGGGCCAGTGTGCGGAAACCCCGGCCAAACTTTAAAGGTGCCATCTG 6600
 6601 GAGGTTGGGGCTCAGAGTACCGGAGGTGACGAGCAAGGTCATACACTACATAC 6660
 6601 GAGGTTGGGGCTCAGAGTACCGGAGGTGACGAGCAAGGTCATACACTACATAC 6660
 6661 AGGACTCACCACTGATACTTGAAGTCCCTGCAACTACCTCTCCCGAGTCTTTTC 6720
 6661 AGGACTCACCACTGATACTTGAAGTCCCTGCAACTACCTCTCCCGAGTCTTTTC 6720
 6721 CTGGGTGGAGCGGAGTGCAGATCCATAGGTTTGCCCCACACCAAGCCGTTTTCCGGGA 6780
 6721 CTGGGTGGAGCGGAGTGCAGATCCATAGGTTTGCCCCACACCAAGCCGTTTTCCGGGA 6780
 6781 TGAGGTCTGTTCTGGTTGGGCTTAATTCATTTGTCGTGGGTCCAGCTTCTTTCGGA 6840
 6781 TGAGGTCTGTTCTGGTTGGGCTTAATTCATTTGTCGTGGGTCCAGCTTCTTTCGGA 6840
 6841 CCTGAAACCCGACACAGAGTATTTGATGTCCTGCTAAAGATCCATCTCATATCACGGC 6900
 6841 CCTGAAACCCGACACAGAGTATTTGATGTCCTGCTAAAGATCCATCTCATATCACGGC 6900
 6901 GGAGACTGACGCGCGGCTTTAGCGGGGGTCAACCCCATCCGAGGCAAGCTCTCGGC 6960
 6901 GGAGACTGACGCGCGGCTTTAGCGGGGGTCAACCCCATCCGAGGCAAGCTCTCGGC 6960
 6961 GAGCCAGTATCGGACCAATCGTGTGGAGCCACTGCAACCCCAAGGCAAGCTATGA 7020
 6961 GAGCCAGTATCGGACCAATCGTGTGGAGCCACTGCAACCCCAAGGCAAGCTATGA 7020
 7021 TGTGACATGTGTGATCTAACTGTTATGCGGGGCGATGACTCGGATAGTCTGG 7080
 7021 TGTGACATGTGTGATCTAACTGTTATGCGGGGCGATGACTCGGATAGTCTGG 7080
 7081 GTCCAAAGTGGTCTGAGTCTCTCGACCCCAATGGTTCGAAGAAAGGAGGACCTTGA 7140
 7081 GTCCAAAGTGGTCTGAGTCTCTCGACCCCAATGGTTCGAAGAAAGGAGGACCTTGA 7140
 7141 GCCTTCGATACCATCAGATATACATGCTCCCAAGAGAGGTTCCCAACAGCTTTACCGGC 7200
 7141 GCCTTCGATACCATCAGATATACATGCTCCCAAGAGAGGTTCCCAACAGCTTTACCGGC 7200
 7201 CTGGGACACGCGCTGATTACAAACCCGCTTGTGGAATCGTGAAGAAAGGCAAGATTACA 7260
 7201 CTGGGACACGCGCTGATTACAAACCCGCTTGTGGAATCGTGAAGAAAGGCAAGATTACA 7260
 7261 ACCGGCCTGTTGGGGGTGTGCTCTCCCTCTCTAGGAAACCCCGACGCTCCCGC 7320
 7261 ACCGGCCTGTTGGGGGTGTGCTCTCCCTCTCTAGGAAACCCCGACGCTCCCGC 7320
 7321 AAGGAGCGCGGACAGTGGGCTTAAGTGAAGGACTCCATAGGAGATGCCCTTCAACAGCT 7380
 7321 AAGGAGCGCGGACAGTGGGCTTAAGTGAAGGACTCCATAGGAGATGCCCTTCAACAGCT 7380
 7381 GGCCATTAAGTCTTTGGCAGCCCGCCCGAGCGGCTTTCAGGCTTTCCACGGGGC 7440
 7381 GGCCATTAAGTCTTTGGCAGCCCGCCCGAGCGGCTTTCAGGCTTTTCACGGGGC 7440
 7441 GGGCGCTGCCGATTCGGGAGTCAAGCGCTCTCTGATGATGTTGGCCCTTTTCGAGACAGG 7500

7441 GGGCGCTGCCGATTCGGGAGTCAAGCGCTCTCTGATGATGTTGGCCCTTTTCGAGACAGG 7500
 7501 TTCCATCTCTTCCATGCCCCCTCGAGGGGAGCTTGGAGATCCAGACCTTGGAGCCTGA 7560
 7501 TTCCATCTCTTCCATGCCCCCTCGAGGGGAGCTTGGAGATCCAGACCTTGGAGCCTGA 7560
 7561 CGAGTGTAGAGCCCAACCCCGCCCGAGGGGGGTGGCAGCTCCCGCTCGGACTCCGG 7620
 7561 CGAGTGTAGAGCCCAACCCCGCCCGAGGGGGGTGGCAGCTCCCGCTCGGACTCCGG 7620
 7621 GTCTGTGTACTTGTCTCCGAGGAGCAGCTCCGTCGTGTGTCTCCATGTCTATCTC 7680
 7621 GTCTGTGTACTTGTCTCCGAGGAGCAGCTCCGTCGTGTGTCTCCATGTCTATCTC 7680
 7681 CTGGACCGGGGCTCTAATACTCTTGTAGTCCCGAGAGGAGGTTTACCGATTAAACC 7740
 7681 CTGGACCGGGGCTCTAATACTCTTGTAGTCCCGAGAGGAGGTTTACCGATTAAACC 7740
 7741 CTGAGCAACTCCCTGTTGCGATATCAACAAGGTGTACTGTACCAACAAGAGAGCGC 7800
 7741 CTGAGCAACTCCCTGTTGCGATATCAACAAGGTGTACTGTACCAACAAGAGAGCGC 7800
 7801 CTCTAAGGGCTAAAAGGTAACTTTTGTAGGATGCAAGTGTCTCTACTACGA 7860
 7801 CTCTAAGGGCTAAAAGGTAACTTTTGTAGGATGCAAGTGTCTCTACTACGA 7860
 7861 CTCTAAGGGCTAAAAGGTAACTTTTGTAGGATGCAAGTGTCTCTACTACGA 7920
 7861 CTCTAAGGGCTAAAAGGTAACTTTTGTAGGATGCAAGTGTCTCTACTACGA 7920
 7921 GGAGAGGCTTTGCGAGTTAACCCCACTTCTGCAAGATCTAAATATGTTGGGGC 7980
 7921 GGAGAGGCTTTGCGAGTTAACCCCACTTCTGCAAGATCTAAATATGTTGGGGC 7980
 7981 TAAGAGGCTCCGAGCTTGTCCGGAGGGCGTTTAAACCAATCAAGTCCGTGTGGAAGA 8040
 7981 TAAGAGGCTCCGAGCTTGTCCGGAGGGCGTTTAAACCAATCAAGTCCGTGTGGAAGA 8040
 8041 CCTCTGAGGACTCAGNAAACCAATTCACCAACCAATATGCGCAAAATGAGTCTT 8100
 8041 CCTCTGAGGACTCAGNAAACCAATTCACCAACCAATATGCGCAAAATGAGTCTT 8100
 8101 CTGCGTGACCCCAAGAGGGGCAAGAAAGAGCTCGCTTATCGTTTACCTGACCT 8160
 8101 CTGCGTGACCCCAAGAGGGGCAAGAAAGAGCTCGCTTATCGTTTACCTGACCT 8160
 8161 CGGCGTCAAGGCTGTGCGAGAAAGATGGCCCTTTATGACATTAACAACAACTTCTCAGGC 8220
 8161 CGGCGTCAAGGCTGTGCGAGAAAGATGGCCCTTTATGACATTAACAACAACTTCTCAGGC 8220
 8221 GGTGATGGGGCTTCTTATGATTCAGATATTCCTCCCGCTCAGCGGGTAGAGTTCTCTT 8280
 8221 GGTGATGGGGCTTCTTATGATTCAGATATTCCTCCCGCTCAGCGGGTAGAGTTCTCTT 8280
 8281 GAAAGCTGGCGGAAAGAGGAGCCCTTATCGGTTTTTCTGATGATACCCGATGTTTGA 8340
 8281 GAAAGCTGGCGGAAAGAGGAGCCCTTATCGGTTTTTCTGATGATACCCGATGTTTGA 8340
 8341 CTCAAACGCTCTAGAGAGAGCATCAGGACTGAGAGGTCCATATATCGGGCTGTCTCTT 8400
 8341 CTCAAACGCTCTAGAGAGAGCATCAGGACTGAGAGGTCCATATATCGGGCTGTCTCTT 8400
 8401 GCCCGAGAGGCGCCACACTGCGCATACACTCGCTAACTGAGAGACTTTTACGTGGAGGGCC 8460
 8401 GCCCGAGAGGCGCCACACTGCGCATACACTCGCTAACTGAGAGACTTTTACGTGGAGGGCC 8460
 8461 TATGTTCAAAGAGGCGCAAACTCGGGGTACAGGGCTTCCCGCGCAGCGGGTGTCT 8520
 8461 TATGTTCAAAGAGGCGCAAACTCGGGGTACAGGGCTTCCCGCGCAGCGGGTGTCT 8520
 8521 CACCACTAGCATGGGGAACCACTCATCATGCTGATGAAAGCTTTAGCGGCTTTGTAAGC 8580

Db	4141	GTCTGGGGGCTACTGTCTAAGGCACATGGCATCAATCCCAACATTAGGACTGGAGTCAG	4200
Qy	4201	GACTGTGACCAACCGGGCGCCCATCAGTACTCTCACATATGGCAAAATCTCTCGCCGATGG	4260
Db	4201	GACTGTGACCAACCGGGCGCCCATCAGTACTCTCACATATGGCAAAATCTCTCGCCGATGG	4260
Qy	4261	GGGCTGTGGGGCGGGCGCCCTACGACATCATATATGTGATGAATGCCCATGCCGTGGACTC	4320
Db	4261	GGGCTGTGGGGCGGGCGCCCTACGACATCATATATGTGATGAATGCCCATGCCGTGGACTC	4320
Qy	4321	TACCACCATCTTTGGCATCGGACAGTCTTGTATCAAGACAGACAGTGGGGTCAGACT	4380
Db	4321	TACCACCATCTTTGGCATCGGACAGTCTTGTATCAAGACAGACAGTGGGGTCAGACT	4380
Qy	4381	AACTGTGCTGGCTACAGCTACGCCCTTGGGTCAAGTCAACACCCCAACCATATAGA	4440
Db	4381	AACTGTGCTGGCTACAGCTACGCCCTTGGGTCAAGTCAACACCCCAACCATATAGA	4440
Qy	4441	GGAGGTGGCCCTTGGGACAGGCGGAGATCCCTTCTATGGAGGGCGATTTCCCTGTGC	4500
Db	4441	GGAGGTGGCCCTTGGGACAGGCGGAGATCCCTTCTATGGAGGGCGATTTCCCTGTGC	4500
Qy	4501	TTACATCAAGGAGGAGACATCTGATCTTCTGCAATTCAGAGAAAGTGTGACGAGCT	4560
Db	4501	TTACATCAAGGAGGAGACATCTGATCTTCTGCAATTCAGAGAAAGTGTGACGAGCT	4560
Qy	4561	CGCGGGCGCCCTTGGGGTATGGGCTTGAATCTGAGTGGCATCTACAGAGGGTTGACAGT	4620
Db	4561	CGCGGGCGCCCTTGGGGTATGGGCTTGAATCTGAGTGGCATCTACAGAGGGTTGACAGT	4620
Qy	4621	CTCCGTAAATACCAACTCAGGAGACGTAGTGGTCTGCCCAACCGACCCCTCATGACAGG	4680
Db	4621	CTCCGTAAATACCAACTCAGGAGACGTAGTGGTCTGCCCAACCGACCCCTCATGACAGG	4680
Qy	4681	GTATATCGGGGACTTTGACTCCGTGATCGATCGATCGATCGATCGATCGATCGATCGAT	4740
Db	4681	GTATATCGGGGACTTTGACTCCGTGATCGATCGATCGATCGATCGATCGATCGATCGAT	4740
Qy	4741	CTTCAGTTTATAGACCCCACTTACCAATACCAACACACAGATTTGCCCTCAAGACGCTGCTC	4800
Db	4741	CTTCAGTTTATAGACCCCACTTACCAATACCAACACACAGATTTGCCCTCAAGACGCTGCTC	4800
Qy	4801	ACGTAGCCAGCGCGGGTGCACGGTAGGGGAAGACTGGGCATTTATAGGTATGTTTC	4860
Db	4801	ACGTAGCCAGCGCGGGTGCACGGTAGGGGAAGACTGGGCATTTATAGGTATGTTTC	4860
Qy	4861	CACGTGTAGCGGAGCGCTCAGGAATGTTTGAAGTGTAGTGTCTGTGAGTGTCTACGACGC	4920
Db	4861	CACGTGTAGCGGAGCGCTCAGGAATGTTTGAAGTGTAGTGTCTGTGAGTGTCTACGACGC	4920
Qy	4921	AGGGCCCGCATGTATGAGCTCACCATCGGAGACCAACCGTCAAGGCTCAGGGCGTATTT	4980
Db	4921	AGGGCCCGCATGTATGAGCTCACCATCGGAGACCAACCGTCAAGGCTCAGGGCGTATTT	4980
Qy	4981	CAACACCGCGGTTTGGCTGTGTGCCAAGACCACTTTGAGTTTGGGAGGCGAGTTTTCAC	5040
Db	4981	CAACACCGCGGTTTGGCTGTGTGCCAAGACCACTTTGAGTTTGGGAGGCGAGTTTTCAC	5040
Qy	5041	CGGCTCACACATAGATGCCCACTTCTTCCCAACCAAGCAATCGGGGGGAAATTTT	5100
Db	5041	CGGCTCACACATAGATGCCCACTTCTTCCCAACCAAGCAATCGGGGGGAAATTTT	5100
Qy	5101	CGCATCTTAAACGCTTACAGGCTACAGTGTGGCTAGGGCCAAAGCCCGCCCGGCTC	5160
Db	5101	CGCATCTTAAACGCTTACAGGCTACAGTGTGGCTAGGGCCAAAGCCCGCCCGGCTC	5160
Qy	5161	CTGGGAGCTCATGTGGAGTGTGTTGACTCGACTCAAGCCCACTCGTGGGCGCCCAACC	5220
Db	5161	CTGGGAGCTCATGTGGAGTGTGTTGACTCGACTCAAGCCCACTCGTGGGCGCCCAACC	5220
Qy	5221	TCTCTGTACCGCTTGGGCTCTGTATCAAGGAGGTCAACCTCATCATCTCCCGTGACAA	5280
Db	5221	TCTCTGTACCGCTTGGGCTCTGTATCAAGGAGGTCAACCTCATCATCTCCCGTGACAA	5280
Qy	5281	ATACATCGCCACCTGATGCAAGCCGACCTTGAGGTCTATGACGACATCGGTCTTGGC	5340
Db	5281	ATACATCGCCACCTGATGCAAGCCGACCTTGAGGTCTATGACGACATCGGTCTTGGC	5340
Qy	5341	AGGGGAGTCTTGGGGCGCTGCGCGGTATTTGCCCTGGCGACCGGGTGTGTTGCATCAT	5400
Db	5341	AGGGGAGTCTTGGGGCGCTGCGCGGTATTTGCCCTGGCGACCGGGTGTGTTGCATCAT	5400
Qy	5401	CGGCGCTTGGCACTTAAACGAGCGAGCGCTGCTTGGCGCGGACCAAGAGGTCTCTATGA	5460
Db	5401	CGGCGCTTGGCACTTAAACGAGCGAGCGCTGCTTGGCGCGGACCAAGAGGTCTCTATGA	5460
Qy	5461	GGCTTTTGTATGATGAGAGGAAATGTGCTCTTAGGGCGGCTCTCATTTGAAGAGGGGAGCG	5520
Db	5461	GGCTTTTGTATGATGAGAGGAAATGTGCTCTTAGGGCGGCTCTCATTTGAAGAGGGGAGCG	5520
Qy	5521	GATAGCGGAGTCTGAACTCCAGATCCAGGCTTATTTGCAGCAGCTTCCAAACAGC	5580
Db	5521	GATAGCGGAGTCTGAACTCCAGATCCAGGCTTATTTGCAGCAGCTTCCAAACAGC	5580
Qy	5581	TCAAGACATACAAACCACTGTGAGGCTTTCATGGGCCCAAGGTAGAAACAATTTCTGGGCCAA	5640
Db	5581	TCAAGACATACAAACCACTGTGAGGCTTTCATGGGCCCAAGGTAGAAACAATTTCTGGGCCAA	5640
Qy	5641	ACACATGTGGAACTTCAATTTAGGGGCTTCAATACCTCGCAGGACTATCAACACTGCCAGG	5700
Db	5641	ACACATGTGGAACTTCAATTTAGGGGCTTCAATACCTCGCAGGACTATCAACACTGCCAGG	5700
Qy	5701	GAACCTGACGATGACTTCAATGATGGCGTTCAGTGGCGGCTCACCAGTCCGCTGCAAC	5760
Db	5701	GAACCTGACGATGACTTCAATGATGGCGTTCAGTGGCGGCTCACCAGTCCGCTGCAAC	5760
Qy	5761	AAGCACCACTATCTCTTCAACATTTTGGGGGCTGCTGAGTATCCCAAAATTCACACACC	5820
Db	5761	AAGCACCACTATCTCTTCAACATTTTGGGGGCTGCTGAGTATCCCAAAATTCACACACC	5820
Qy	5821	CGCGGGGCGCACTGGCTTCTGATGCGCTTGTAGTGGGAGTCCGCTAGGCAAGTATAGG	5880
Db	5821	CGCGGGGCGCACTGGCTTCTGATGCGCTTGTAGTGGGAGTCCGCTAGGCAAGTATAGG	5880
Qy	5881	CTTAGTAAAGTGTAGTGGACATCTTGGCAGGGTATGGTGGGGCTTTCGGGGGCTCT	5940
Db	5881	CTTAGTAAAGTGTAGTGGACATCTTGGCAGGGTATGGTGGGGCTTTCGGGGGCTCT	5940
Qy	5941	CGTGCATTTCAAGATCATGTCTGGCGAAGAGCCCTCCATGGAGGATGTCTCAACTTGTCT	6000
Db	5941	CGTGCATTTCAAGATCATGTCTGGCGAAGAGCCCTCCATGGAGGATGTCTCAACTTGTCT	6000
Qy	6001	GCCTGAAATTTCTGCTCCGGGTGCTTGTAGTGGAGTCACTGCGGGGCTCATTTCTGCG	6060
Db	6001	GCCTGAAATTTCTGCTCCGGGTGCTTGTAGTGGAGTCACTGCGGGGCTCATTTCTGCG	6060
Qy	6061	CGGACAGTGGGACCGGGGGAGGCGCGTCCATGAGTGAATAGACTCATTTCCGCTTTC	6120
Db	6061	CGGACAGTGGGACCGGGGGAGGCGCGTCCATGAGTGAATAGACTCATTTCCGCTTTC	6120
Qy	6121	TTCCAGAGGAAATCAGTCTGCGCCCAACCACTACGTGACGAGTCCGATGCGTCCGAGCG	6180
Db	6121	TTCCAGAGGAAATCAGTCTGCGCCCAACCACTACGTGACGAGTCCGATGCGTCCGAGCG	6180
Qy	6181	TGTGACCCCAACTTCTGGCTCCCTTACCATTAACCAAGCTGCTCAGAGAGTCCCAAACTG	6240
Db	6181	TGTGACCCCAACTTCTGGCTCCCTTACCATTAACCAAGCTGCTCAGAGAGTCCCAAACTG	6240
Qy	6241	GATTACTGAGGACTGCGGCTCCCATGCGGCGGCTCGTGGCTCCGCGATGTGTGGGACTG	6300
Db	6241	GATTACTGAGGACTGCGGCTCCCATGCGGCGGCTCGTGGCTCCGCGATGTGTGGGACTG	6300
Qy	6301	GGTTTCCACCATCTTAACAGACTTTTAAATAATTTGGCTGACTCCAAATTTATTTCCAAAGAT	6360
Db	6301	GGTTTCCACCATCTTAACAGACTTTTAAATAATTTGGCTGACTCCAAATTTATTTCCAAAGAT	6360

3061	AGGTGCTTTGACGCGGTGCCGTACTTCGTGAGGGCTCACGCTCTACTTGAGATGTGCAC	3120	
3121	CATGCAAGGCATCTCGCGGGGGCAGGTACGTCACAGATGGCGCTACTAGCCCTTTGGCAG	3180	
3121	CATGCAAGGCATCTCGCGGGGGCAGGTACGTCACAGATGGCGCTACTAGCCCTTTGGCAG	3180	
3181	GTGAGCTGGCACTTACATCTATGACACCACTCACCCCTATGTCCGATTTGGGCTGTAGTGG	3240	
3181	GTGAGCTGGCACTTACATCTATGACACCACTCACCCCTATGTCCGATTTGGGCTGTAGTGG	3240	
3241	CCTCGGGACCTGCGCGTTCGGCTTACGCTTATCATCTTCAGTCGATGGAGAGAAAGT	3300	
3241	CCTCGGGACCTGCGCGTTCGGCTTACGCTTATCATCTTCAGTCGATGGAGAGAAAGT	3300	
3301	CATTGCTCGGGAGCGGAGACAGCTGCTTGTGGGGACATTTTACACGACATTCCTCGTGTC	3360	
3301	CATTGCTCGGGAGCGGAGACAGCTGCTTGTGGGGACATTTTACACGACATTCCTCGTGTC	3360	
3361	CGCCCGACTTGGTTCGGGAGTCTCTCTTGGGCCACGTGATGGCTATACCTTCAAGGGGTG	3420	
3361	CGCCCGACTTGGTTCGGGAGTCTCTCTTGGGCCACGTGATGGCTATACCTTCAAGGGGTG	3420	
3421	GAGTCTTCTCGCCCCCATCACTGCTTACGCCCCAGGACACGTTGGCTTTTGGGCAACCAT	3480	
3421	GAGTCTTCTCGCCCCCATCACTGCTTACGCCCCAGGACACGTTGGCTTTTGGGCAACCAT	3480	
3481	AGTGTGAGCATGACGGGGCGGCAAGACAGAACAGGCTCGGGAAATTCAGTCTCTGTC	3540	
3481	AGTGTGAGCATGACGGGGCGGCAAGACAGAACAGGCTCGGGAAATTCAGTCTCTGTC	3540	
3541	CACAGTCACTCAGTCTCTTCCTCGAAACATCCATCTCGGGGGTTTTGTGGACTGTCTACCA	3600	
3541	CACAGTCACTCAGTCTCTTCCTCGAAACATCCATCTCGGGGGTTTTGTGGACTGTCTACCA	3600	
3601	TGGAGCTGGCAACAGACTCTGGCCGGCTCAGGGGTCCTCGGACGAGATGTACTCCAG	3660	
3601	TGGAGCTGGCAACAGACTCTGGCCGGCTCAGGGGTCCTCGGACGAGATGTACTCCAG	3660	
3661	TGCTAGGGGACATTAGTAGGGTGGCCAGCCGCCCTGGGACTAAATCTTTGGAGCCGTG	3720	
3661	TGCTAGGGGACATTAGTAGGGTGGCCAGCCGCCCTGGGACTAAATCTTTGGAGCCGTG	3720	
3721	CACGTGTGGAGCGGTGCACTGTACCTGGTCACGCGGAACGCTGATGTCATCCCGGCTCG	3780	
3721	CACGTGTGGAGCGGTGCACTGTACCTGGTCACGCGGAACGCTGATGTCATCCCGGCTCG	3780	
3781	AAGACGGGGGACAAACGGGAGGCGCTACTCTCCCGAGACCTCTTTCACCTTGAAGGG	3840	
3781	AAGACGGGGGACAAACGGGAGGCGCTACTCTCCCGAGACCTCTTTCACCTTGAAGGG	3840	
3841	GTCCTCAGGAGCCGGTCTATGCCCAGGGGCCACGCTGCGGAGTCTTCCCGGGCAGC	3900	
3841	GTCCTCAGGAGCCGGTCTATGCCCAGGGGCCACGCTGCGGAGTCTTCCCGGGCAGC	3900	
3901	TGTGTGCTCTCGGGCGTGGCTAAAGTCATAGATTTTCATCCCGCTTGAGACACTCGACAT	3960	
3901	TGTGTGCTCTCGGGCGTGGCTAAAGTCATAGATTTTCATCCCGCTTGAGACACTCGACAT	3960	
3961	CGTCAACGGGTCCCCCACTTTTGTGACAAACAGCACCACTGCTGTGTCGCCAGACCTTA	4020	
3961	CGTCAACGGGTCCCCCACTTTTGTGACAAACAGCACCACTGCTGTGTCGCCAGACCTTA	4020	
4021	TCAGGTGGGTACTTTGCATGCCCGACTGGCAGTGGAAAGAGCAACCAAGTTCCTGTGCG	4080	
4021	TCAGGTGGGTACTTTGCATGCCCGACTGGCAGTGGAAAGAGCAACCAAGTTCCTGTGCG	4080	
4081	ATATGCTGCTCAGGGGTATAAAGTGTAGTCTTAATCCCTCAGTGGCTGCCACCTGGG	4140	
4081	ATATGCTGCTCAGGGGTATAAAGTGTAGTCTTAATCCCTCAGTGGCTGCCACCTGGG	4140	
4141	GTTTGGGCGTACTTGTCTAAGGACATGGGCATCAATCCCAACATTAGGACTGGAGTCAG	4200	

JOURNAL Patent: WO 00/5338-A 1 14-DEC-2000;
THE GOVERNMENT OF THE UNITED STATES OF AMERICA (US)

FEATURES
source Location/Qualifiers

1..9711
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
BASE COUNT 1987 a 2853 c 2668 g 2203 t
ORIGIN

Query Match 100.0%; Score 9711; DB 6; Length 9711;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 9711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACCCGCCCTTAATAGGGGCGACACTCGGCATCAATCACTCCCTCTGAGGAACTACTGT 60
DB 1 ACCCGCCCTTAATAGGGGCGACACTCGGCATCAATCACTCCCTCTGAGGAACTACTGT 60
QY 61 CTTTACGCGAAGAGCGCTTAGCCATCGCGTTAGTATGAGTGTGCTACAGCCTCCAGGCC 120
DB 61 CTTTACGCGAAGAGCGCTTAGCCATCGCGTTAGTATGAGTGTGCTACAGCCTCCAGGCC 120
QY 121 CCCCTCCCGGAGAGCGATAGTGGTCTCGGGAACCGGTGAGTACACCGGAATTGCGGG 180
DB 121 CCCCTCCCGGAGAGCGATAGTGGTCTCGGGAACCGGTGAGTACACCGGAATTGCGGG 180
QY 181 AAGACTGGGTCTTCTTGGATAAACCACTATGCGCGGCCATTTGGGCGTGCCTCCG 240
DB 181 AAGACTGGGTCTTCTTGGATAAACCACTATGCGCGGCCATTTGGGCGTGCCTCCG 240
QY 241 CAAGACTGCTAGCGGAGTAGCGTTGGGTTGCGAAGGCGCTTGGTACTGCTGTATAGGG 300
DB 241 CAAGACTGCTAGCGGAGTAGCGTTGGGTTGCGAAGGCGCTTGGTACTGCTGTATAGGG 300
QY 301 TGCTTGGAGTGGCCCGGAGGTCTCGTAGACCGTGCACCATGAGCACAATCTTAACC 360
DB 301 TGCTTGGAGTGGCCCGGAGGTCTCGTAGACCGTGCACCATGAGCACAATCTTAACC 360
QY 361 TCAAGAAACCAAGAAACCAACCACTGCGCCCAAGAGCTTAAAGTTTCCGGCGG 420
DB 361 TCAAGAAACCAAGAAACCAACCACTGCGCCCAAGAGCTTAAAGTTTCCGGCGG 420
QY 421 CGGCCAGATCGTTGGCGGAGTATCTTGTGTCGCGCAGGCGGCCCAAGGTTGGGTGCG 480
DB 421 CGGCCAGATCGTTGGCGGAGTATCTTGTGTCGCGCAGGCGGCCCAAGGTTGGGTGCG 480
QY 481 CGCGAAGAGAGACTTCGAGGCGTCCAGCAGCAGTGGAGGCGCCAGCCATCCCTAA 540
DB 481 CGCGAAGAGAGACTTCGAGGCGTCCAGCAGCAGTGGAGGCGCCAGCCATCCCTAA 540
QY 541 AGATCGGCGTCCACTGCGCAATCTCGGGAACCAAGGATACCCCTGGCGCCCTATACG 600
DB 541 AGATCGGCGTCCACTGCGCAATCTCGGGAACCAAGGATACCCCTGGCGCCCTATACG 600
QY 601 GAATGAGGAGTCTCGGCTGGCAGAGTCTCTGTCCTCCCGAGGTTCCCGTCCCTTTG 660
DB 601 GAATGAGGAGTCTCGGCTGGCAGAGTCTCTGTCCTCCCGAGGTTCCCGTCCCTTTG 660
QY 661 GGGCCCAATGACCCCGGATAGGTTCGCGCAAGCTGGGTAAAGTCAATGATACCCCTAAC 720
DB 661 GGGCCCAATGACCCCGGATAGGTTCGCGCAAGCTGGGTAAAGTCAATGATACCCCTAAC 720
QY 721 GTGCGGCTTTGCGGACTCATGGGTATACCTCTGTCGTCGCGCCCGCTCGCGCGCGT 780
DB 721 GTGCGGCTTTGCGGACTCATGGGTATACCTCTGTCGTCGCGCCCGCTCGCGCGCGT 780
QY 781 CGCCAGAGCTCTCGCGATGGCGTGGAGTCTCTGAGGAGCGGGTTAAATTTGCAACAGG 840
DB 781 CGCCAGAGCTCTCGCGATGGCGTGGAGTCTCTGAGGAGCGGGTTAAATTTGCAACAGG 840
QY 841 GAATTTACCGGTTGCTCTCTTTCTATCTTCTGCTGCGCCCTGCTGCTGATCACCAC 900
DB 841 GAATTTACCGGTTGCTCTCTTTCTATCTTCTGCTGCGCCCTGCTGCTGATCACCAC 900

QY 901 CCCGGTCTCCGCTGCCGAAGTGAAGAACATCACTAGTACCGCTACATGGTGACTAACGACTG 960
DB 901 CCCGGTCTCCGCTGCCGAAGTGAAGAACATCACTAGTACCGCTACATGGTGACTAACGACTG 960
QY 961 CACCAATGACAGCATTAATCTGCGAGCTCCAGGCTGCTCTCCAGCTCCCGGGTGGCT 1020
DB 961 CACCAATGACAGCATTAATCTGCGAGCTCCAGGCTGCTCTCTCCAGCTCCCGGGTGGCT 1020
QY 1021 CCCGTGCGAGAAAGTGGGAATGCAATCTCAGTGTCTGATACCGTCTCAACCAATGTGGC 1080
DB 1021 CCCGTGCGAGAAAGTGGGAATGCAATCTCAGTGTCTGATACCGTCTCAACCAATGTGGC 1080
QY 1081 CGTGCAGCGCGCGCGCCCTCAACGAGGCTTGGCGACGCAATGACATGATGTTGTGAT 1140
DB 1081 CGTGCAGCGCGCGCGCCCTCAACGAGGCTTGGCGACGCAATGACATGATGTTGTGAT 1140
QY 1141 GTCCGCGACGCTCTGCTCTGCTCTCTACGTGGGGGACCTCTCGGTGGGGTGTGCTCGC 1200
DB 1141 GTCCGCGACGCTCTGCTCTGCTCTCTACGTGGGGGACCTCTCGGTGGGGTGTGCTCGC 1200
QY 1201 AGCCCAATGTTTCAATGTTCTCGCGCAGCACCACTGGTGTGTTGTTCCAAAGACTGCAATTTGCTC 1260
DB 1201 AGCCCAATGTTTCAATGTTCTCGCGCAGCACCACTGGTGTGTTGTTCCAAAGACTGCAATTTGCTC 1260
QY 1261 CATCTACCTTGTGTACCATCACTGGACACCGCATGGCATGGGACATGATGAATGCTGTC 1320
DB 1261 CATCTACCTTGTGTACCATCACTGGACACCGCATGGCATGGGACATGATGAATGCTGTC 1320
QY 1321 GCGCAGCGGTACGATGATCTTGGCGTACCGGATGCGGTGTCGCGAGTCAATATACAT 1380
DB 1321 GCGCAGCGGTACGATGATCTTGGCGTACCGGATGCGGTGTCGCGAGTCAATATACAT 1380
QY 1381 CATTAGCGGGCTCATTTGGGGCTCATGTTTGGCTTGGCTTACTTCTTATGCAAGGAGC 1440
DB 1381 CATTAGCGGGCTCATTTGGGGCTCATGTTTGGCTTGGCTTACTTCTTATGCAAGGAGC 1440
QY 1441 GTGGGCGAAAGTCTGTTGTCTCTTGTGGCGCGCGGGTGGACGCGCGACCCATAC 1500
DB 1441 GTGGGCGAAAGTCTGTTGTCTCTTGTGGCGCGCGGGTGGACGCGCGACCCATAC 1500
QY 1501 TGTGGGGTCTTGGCGCGCAGACCAACCGGGGCTTCCAGCTTATTTGACATGGGCCC 1560
DB 1501 TGTGGGGTCTTGGCGCGCAGACCAACCGGGGCTTCCAGCTTATTTGACATGGGCCC 1560
QY 1561 CAGGCGAGAAATCCAGCTCGTTAAACCAATGCGAGCTGGCACATCAACCGCACCGCCCT 1620
DB 1561 CAGGCGAGAAATCCAGCTCGTTAAACCAATGCGAGCTGGCACATCAACCGCACCGCCCT 1620
QY 1621 GAACTGCAATGACTCTCTTGCACACCGGCTTTATCGGCTCTCTTCTTACACCCACAGCTT 1680
DB 1621 GAACTGCAATGACTCTCTTGCACACCGGCTTTATCGGCTCTCTTCTTACACCCACAGCTT 1680
QY 1681 CAATCTGTCAGATGTCGCGAAGCATGTCGCGCTGCGCGAGTATCGAGGCTTCCGGGT 1740
DB 1681 CAATCTGTCAGATGTCGCGAAGCATGTCGCGCTGCGCGAGTATCGAGGCTTCCGGGT 1740
QY 1741 GGGATGGGCGCTTGCATATGAGGATATGAGGATATGTCACCAATCCAGAGGATGAGACCTTA 1800
DB 1741 GGGATGGGCGCTTGCATATGAGGATATGAGGATATGTCACCAATCCAGAGGATGAGACCTTA 1800
QY 1801 TTGCTGGCACTACCCCAAGGAGTGTGCGGTGCTTCCGCAAGACTGTGTGTGGCCC 1860
DB 1801 TTGCTGGCACTACCCCAAGGAGTGTGCGGTGCTTCCGCAAGACTGTGTGTGGCCC 1860
QY 1861 AGTGTACTTTCACCCCGAGCCAGTGTGTGAGGAGCAGCAGGCTTTCGAGGCGC 1920
DB 1861 AGTGTACTTTCACCCCGAGCCAGTGTGTGAGGAGCAGCAGGCTTTCGAGGCGC 1920
QY 1921 CACTTACAGTGGGGGAGAAATGAGACAGATGTTCTTCTTATGAAAGCACTCGACCC 1980
DB 1921 CACTTACAGTGGGGGAGAAATGAGACAGATGTTCTTCTTATGAAAGCACTCGACCC 1980

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: February 12, 2004, 06:04:40 ; Search time 33743 Seconds
(without alignments)
11773.505 Million cell updates/sec
Title: US-09-980-559-1
Perfect score: 9711
Sequence: 1 accgcccttaataaggcg.....gggtctctcgagcatcatgt 9711

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 2888711 seqs, 2045481386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenBank:

- 1: gb_ba.*
- 2: gb_hg.*
- 3: gb_in.*
- 4: gb_on.*
- 5: gb_ov.*
- 6: gb_ph.*
- 7: gb_pl.*
- 8: gb_pr.*
- 9: gb_ro.*
- 10: gb_ro.*
- 11: gb_ro.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_on.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_scs.*
- 28: em_un.*
- 29: em_vl.*
- 30: em_hg_hum.*
- 31: em_hg_inv.*
- 32: em_hg_other.*
- 33: em_hg_mus.*
- 34: em_hg_pln.*
- 35: em_hg_rod.*
- 36: em_hg_mam.*
- 37: em_hg_vrt.*
- 38: em_sv.*
- 39: em_hgo_hum.*
- 40: em_hgo_mus.*
- 41: em_hgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	9711	100.0	9711	6	AX057086	AX057086 Sequence
2	9711	100.0	9711	6	AX057317	AX057317 Sequence
3	9711	100.0	9711	6	AX057395	AX057395 Sequence
4	8973.6	92.4	9589	14	AF177036	AF177036 Hepatitis C
5	8970.4	92.4	9589	6	HPCPOLP	D00944 Hepatitis C
6	8970.4	92.4	9589	6	E07361	E07361 gRNA of Hep
7	8970.4	92.4	9589	6	E07362	E07362 cDNA of Hep
8	8970.4	92.4	9589	6	I12861	I12861 Sequence 2
9	8507.4	87.6	9659	14	AB047641	AB047641 Hepatitis
10	8425.8	86.2	9693	14	AF169005	AF169005 Hepatitis
11	8370.2	86.2	9693	14	AF169003	AF169003 Hepatitis
12	8367.8	86.2	9653	14	AF169004	AF169004 Hepatitis
13	8364.2	86.1	9683	14	AB047645	AB047645 Hepatitis
14	8349	86.0	9691	14	AB047644	AB047644 Hepatitis
15	8329.4	85.8	9661	14	AF169002	AF169002 Hepatitis
16	8273.4	85.2	9674	14	AB047640	AB047640 Hepatitis
17	8233.4	84.8	9678	14	AB047642	AB047642 Hepatitis
18	8210.2	84.5	9416	14	AF238481	AF238481 Hepatitis
19	8186.8	84.3	9416	14	AF238482	AF238482 Hepatitis
20	8169.2	84.1	9416	14	AF238483	AF238483 Hepatitis
21	8156.4	84.0	9416	14	AF238484	AF238484 Hepatitis
22	8126.4	83.7	9416	14	AF238485	AF238485 Hepatitis
23	8083.4	83.2	9641	14	AB047643	AB047643 Hepatitis
24	8020.6	82.6	9678	6	BD160781	BD160781 Gene of f
25	8020.6	82.6	9678	14	AB047639	AB047639 Hepatitis
26	6479.2	65.7	9488	14	AB031663	AB031663 Hepatitis C
27	6383.6	65.7	9488	14	AB030907	AB030907 Hepatitis
28	6048	62.3	9654	14	AB030907	AB030907 Hepatitis
29	6031.4	62.1	9511	6	E07468	E07468 cDNA sequen
30	6031.4	62.1	9511	6	I12865	I12865 Sequence 7
31	6027	62.1	9511	6	E07467	E07467 gDNA sequen
32	5993.8	61.7	9511	14	HPCJ8G	D00988 Hepatitis C
33	5936	61.1	9416	14	AF238486	AF238486 Hepatitis
34	5816.8	59.9	9611	14	AF177037	AF177037 Hepatitis
35	5800.8	59.7	9611	6	AX057088	AX057088 Sequence
36	5800.8	59.7	9611	6	AX057092	AX057092 Sequence
37	5800.8	59.7	9611	14	AF177039	AF177039 Hepatitis
38	5692	58.6	9611	6	AX057090	AX057090 Sequence
39	5692	58.6	9611	14	AF177038	AF177038 Hepatitis
40	5676	58.4	9611	6	AX057094	AX057094 Sequence
41	5676	58.4	9611	14	AF177040	AF177040 Hepatitis
42	4625.6	47.6	9646	6	AR110828	AR110828 Sequence
43	4625.6	47.6	9646	6	BD069982	BD069982 Functiona
44	4625.6	47.6	9646	14	AF009606	AF009606 Hepatitis
45	4620.8	47.6	9646	12	AF387806	AF387806 Synthetic

ALIGNMENTS

RESULT 1
AX057086
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE

AX057086
Sequence 1 from Patent WO0075338.
AX057086
AX057086.1 GI:12309919

Hepatitis C virus
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.

Yanagi,M., Bukh,J., Emerson,S.U. and Purcell,R.H.
Cloned genome of infectious hepatitis c virus of genotype 2a and
uses thereof

```
=> "HCV genotype 2a isolate"
    7349 "HCV"
    17 "HCVS"
    7353 "HCV"
        ("HCV" OR "HCVS")
    30644 "GENOTYPE"
    39594 "GENOTYPES"
    53369 "GENOTYPE"
        ("GENOTYPE" OR "GENOTYPES")
    27493 "2A"
    49972 "ISOLATE"
    58823 "ISOLATES"
    98536 "ISOLATE"
        ("ISOLATE" OR "ISOLATES")
L1      0 "HCV GENOTYPE 2A ISOLATE"
        ("HCV" (W) "GENOTYPE" (W) "2A" (W) "ISOLATE")
```

```
=> "HCV genotype 2a"
    7349 "HCV"
    17 "HCVS"
    7353 "HCV"
        ("HCV" OR "HCVS")
    30644 "GENOTYPE"
    39594 "GENOTYPES"
    53369 "GENOTYPE"
        ("GENOTYPE" OR "GENOTYPES")
    27493 "2A"
L2      15 "HCV GENOTYPE 2A"
        ("HCV" (W) "GENOTYPE" (W) "2A")
```

```
=> L2 and "cell culture"
    1743376 "CELL"
    1554048 "CELLS"
    2339357 "CELL"
        ("CELL" OR "CELLS")
    343995 "CULTURE"
    191184 "CULTURES"
    456885 "CULTURE"
        ("CULTURE" OR "CULTURES")
    61464 "CELL CULTURE"
        ("CELL" (W) "CULTURE")
L3      0 L2 AND "CELL CULTURE"
```

```
=> DIS L2 1- IBIB IABS
YOU HAVE REQUESTED DATA FROM 15 ANSWERS - CONTINUE? Y/(N):Y
THE ESTIMATED COST FOR THIS REQUEST IS 38.12 U.S. DOLLARS
DO YOU WANT TO CONTINUE WITH THIS REQUEST? (Y)/N:Y
```

```
L2  ANSWER 1 OF 15  CAPLUS  COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER:      2003:463142  CAPLUS
DOCUMENT NUMBER:      139:229074
TITLE:                Sequence analysis of PePHD within HCV E2 region and
                        correlation with resistance of interferon therapy in
                        Japanese patients infected with HCV
                        genotypes 2a and 2b
AUTHOR(S):            Saito, Takeshi; Ito, Takayoshi; Ishiko, Hiroaki;
                        Yonaha, Mari; Morikawa, Kenichi; Miyokawa, Akio;
                        Mitamura, Keiji
CORPORATE SOURCE:      Second Department of Internal Medicine, Showa
                        University School of Medicine, Tokyo, Japan
SOURCE:                American Journal of Gastroenterology (2003), 98(6),
                        1377-1383
                        CODEN: AJGAAR; ISSN: 0002-9270
PUBLISHER:             Elsevier Science Inc.
```


DOCUMENT TYPE: Journal
LANGUAGE: English
ABSTRACT:

Objective. Hepatitis C virus (HCV) E2 protein was recently reported to have a double-stranded RNA-activated protein kinase-eukaryotic initiation factor 2 α (PKR-eIF2 α) phosphorylation homol. domain (PePHD); PKR is induced by interferon (IFN). PePHD interacts with PKR and inactivates it. PePHD could be a predictor for IFN response, like the interferon sensitivity determination region (ISDR) of HCV NS5A. Several groups reported that PePHD is conserved, and mutations in this region do not correlate with IFN response. In this study, we further investigated the amino acid variation of PePHD among four major genotypes and its correlation with IFN response. Methods. We enrolled 74 patients for this study and determined PePHD sequence of HCV derived from sera of patients infected with HCV genotype 1a (1 patient; nonresponder [NR]), 1b (36 patients; 4 complete responders [CR], 32 NR), 2a (29 patients; 17 CR, 12 NR), and 2b (8 patients; 3 CR, 5 NR). We also analyzed mutations in ISDR of HCV genotype 1b in 31 patients. Results. PePHD had several variations among four genotypes investigated. In patients infected with HCV genotype 1b, PePHD sequence was well conserved and seemed to have no correlation with IFN response. Mutations in ISDR were correlated with IFN response. In patients with **HCV genotypes 2a** and 2b, PePHD had multiple variations, and one particular motif, "RGQQ-" at the N-terminus, showed a close correlation with IFN resistance. All eight patients with HCV containing this motif were IFN nonresponders. Conclusions. IFN resistance of HCV correlates with its "RGQQ-" motif at the N-terminus of PePHD in **HCV genotype ***2a***** and 2b. PePHD of HCV could be a predictor of IFN resistance in patients infected with **HCV genotype 2a** and 2b.

REFERENCE COUNT: 34 THERE ARE 34 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L2 ANSWER 2 OF 15 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 2003:153966 CAPLUS

DOCUMENT NUMBER: 138:319442

TITLE: Association of amino acid substitution pattern in nonstructural protein 5A of hepatitis C virus genotype2a low viral load and response to interferon monotherapy

AUTHOR(S): Akuta, Norio; Suzuki, Fumitaka; Tsubota, Akihito; Suzuki, Yoshiyuki; Hosaka, Tetsuya; Someya, Takashi; Kobayashi, Masahiro; Saitoh, Satoshi; Arase, Yasuji; Ikeda, Kenji; Kumada, Hiromitsu

CORPORATE SOURCE: Department of Gastroenterology, Toranomon Hospital, Tokyo, 105-0001, Japan

SOURCE: Journal of Medical Virology (2003), 69(3), 376-383
CODEN: JMVIDB; ISSN: 0146-6615

PUBLISHER: Wiley-Liss, Inc.

DOCUMENT TYPE: Journal

LANGUAGE: English

ABSTRACT:

Patients with low titer (<0.5 mEq/mL) of hepatitis C virus (HCV) *****genotype*** 2a** achieve high and sustained response (SR) rates to interferon (IFN) monotherapy, but we also encounter patients who are resistant to therapy. We explored the relationship between response to IFN and virol. differences in such patients. We evaluated 159 consecutive naive patients with low titer of **HCV genotype 2a** who received IFN monotherapy. A case-control study matched for age, sex, and viral load was conducted to examine the substitution patterns in amino acid positions (amino acids) 2163-2254 of nonstructural (NS) 5A between nonresponders to ideal IFN dose (≥ 500 million units) (nonresponders; NR) and responder to less than ideal dose. Overall, 82.4% achieved SR. The substitution nos. in amino acids 2193-2254 were higher in SR than NR patients ($P < 0.05$). High proportions of patients with substitution at amino acid 2205 (mainly threonine [T] instead of alanine [A]), dual amino acids 2169 and 2205 (mainly A-T instead of T-A), and

those without substitution at amino acids 2227 were NR ($P < 0.05$). Four of 7 NR patients achieved SR after receiving a second course of IFN. Their amino acids patterns at positions probably associated with sensitivity to IFN did not change at the start of initial and second therapies except for one patient, and they had lower viral load and were treated with higher IFN dose in the second course compared with the initial course. Our results suggest that substitution patterns in NS5A in patients with low titer of **HCV genotype**

2a may affect their response to IFN, but the response to therapy may be affected by mechanisms other than substitutions in this region.

REFERENCE COUNT: 40 THERE ARE 40 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L2 ANSWER 3 OF 15 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 2003:153965 CAPLUS

DOCUMENT NUMBER: 138:396946

TITLE: Hepatitis C virus molecular epidemiology in Uzbekistan

AUTHOR(S): Kurbanov, Fuat; Tanaka, Yasuhito; Sugauchi, Fuminaka; Kato, Hideaki; Ruzibakiev, Ruslan; Zalyalieva, Mariam; Yunusova, Zulhumor; Mizokami, Masashi

CORPORATE SOURCE: Department of Clinical Molecular Informative Medicine, Nagoya City University Graduate School of Medical Sciences, Nagoya, 467-8601, Japan

SOURCE: Journal of Medical Virology (2003), 69(3), 367-375

CODEN: JMVIDB; ISSN: 0146-6615

PUBLISHER: Wiley-Liss, Inc.

DOCUMENT TYPE: Journal

LANGUAGE: English

ABSTRACT:

The aim of this study was to identify hepatitis C virus (HCV) genotypes and to estimate their prevalence in various risk groups and the regional distribution in Uzbekistan. Preliminary serol. screening of 1,269 subjects revealed 6.5% anti-HCV-pos. in a general population, 27.1% in patient groups, and 51.7% among i.v. drug users. HCV genotypes of 104 anti-HCV-pos. subjects were determined using a PCR-genotyping system in core region, and the results were supported by nucleotide sequencing of the NS5B region. Genotype 1b identified in total 64.2%, was the most prevalent. The genotype 3a identified in 25.0% was the second one distributed. **HCV genotypes 2a, 1a, 2b, and 3b** were identified in 3.8%, 2.9%, 2.9%, and 1.0% of cases, resp. The i.v. drug users were distinguished from other groups by having the highest prevalence of genotype 3a, i.e., 50.0%, higher than the 33.3% for genotype 1b in this group. Geog., genotype 1b was common; genotype 3a was also found frequently in all three regions. Uncommon HCV genotypes (1a, 2a, 2b, and 3b) were found in comparatively greater variability in the western region. Mol. evolutionary anal. based on the NS5B region did not reveal specific clustering or indigenous strains among Uzbekistan HCV isolates. In summary, two main mechanisms of HCV infection distribution were observed in Uzbekistan: HCV 1b genotype infection is wide-spread through blood products, and HCV 3a genotype infection is spreading through the growing number of i.v. drug users.

REFERENCE COUNT: 30 THERE ARE 30 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L2 ANSWER 4 OF 15 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 2003:136375 CAPLUS

DOCUMENT NUMBER: 138:280774

TITLE: SEN virus infection in patients with chronic hepatitis C: preferential coinfection with hepatitis C genotype 2a and no effect on response to therapy with interferon plus ribavirin

AUTHOR(S): Kao, Jia-Horng; Chen, Wendy; Chen, Pei-Jer; Lai, Ming-Yang; Chen, Ding-Shinn

CORPORATE SOURCE: Graduate Institute of Clinical Medicine, Department of Internal Medicine, National Taiwan University College

of Medicine and National Taiwan University Hospital,
Taipei, Taiwan
SOURCE: Journal of Infectious Diseases (2003), 187(2), 307-310
CODEN: JIDIAQ; ISSN: 0022-1899
PUBLISHER: University of Chicago Press
DOCUMENT TYPE: Journal
LANGUAGE: English

ABSTRACT:

To clarify the influence that a recently identified SEN virus (SENV) has on hepatitis C virus (HCV) response to therapy with interferon plus ribavirin, 2 SENV variants, SENV-D and SENV-H, were studied in 100 patients with chronic hepatitis C; 57 of these patients were pos. for SENV-D/H DNA, and there were no differences, in clinicopathol. features, between patients with and without SENV coinfection. However, patients with SENV coinfection had a higher prevalence of **HCV genotype 2a** than did those without it.

The sustained HCV response rate after combination therapy was comparable between patients with and without SENV coinfection. Of the 57 patients with SENV coinfection, 18 (32%) had a sustained SENV response to combination therapy, and SENV-D had a higher sustained response rate than did SENV-H.

These results suggest that SENV has a specific link to **HCV**

*****genotype*** 2a** and that SENV infection has no apparent effect on coexisting chronic hepatitis C.

REFERENCE COUNT: 15 THERE ARE 15 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L2 ANSWER 5 OF 15 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 2002:874075 CAPLUS

DOCUMENT NUMBER: 138:168479

TITLE: Efficacy of interferon monotherapy to 394 consecutive naive cases infected with hepatitis C virus genotype 2a in Japan: therapy efficacy as consequence of tripartite interaction of viral, host and interferon treatment-related factors

AUTHOR(S): Akuta, Norio; Suzuki, Fumitaka; Tsubota, Akihito; Suzuki, Yoshiyuki; Someya, Takashi; Kobayashi, Masahiro; Saitoh, Satoshi; Arase, Yasuji; Ikeda, Kenji; Kumada, Hiromitsu

CORPORATE SOURCE: Division of Gastroenterology, Toranomon Hospital, Minato-ku, Tokyo, 105-0001, Japan

SOURCE: Journal of Hepatology (2002), 37(6), 831-836

CODEN: JOHEEC; ISSN: 0168-8278

PUBLISHER: Elsevier Science Ltd.

DOCUMENT TYPE: Journal

LANGUAGE: English

ABSTRACT:

Background/Aims: The mechanism of variable response to interferon (IFN) monotherapy in patients infected with **HCV genotype**

*****2a***** is still unclear. Here we investigated the response in a large

group of patients infected with genotype 2a. Methods: We evaluated 394 consecutive non-cirrhotic naive patients infected with genotype 2a who received IFN monotherapy for 24 wk, including initial aggressive induction therapy. Of these, 97 were also evaluated for early viral kinetics in serum and treatment efficacy. Results: The overall sustained response (SR) rate was 68.3% (viral load <1.0 Meq/mL (82.4%); ≥1.0 (52.4%)). Multivariate anal. identified five independent factors associated with SR; viral load <1.0 Meq/mL, total IFN dose ≥700 million units, hepatocyte steatosis none or mild, albumin ≥3.9 g/dL, and alanine aminotransferase ≥75 IU/l. The kinetic study showed that serum viral clearance at ≤1 wk was the best predictor of SR, and persistence at ≥ 4 wk was a predictor of non-SR.

Conclusions: Our study suggests that viral, host and IFN treatment-related factors determine the response to IFN monotherapy in patients infected with

*****HCV*** genotype 2a**. Further, we report that IFN

monotherapy is very effective for patients with genotype 2a, especially for those

with low viral load; and that early viral kinetics is useful as a predictor of the response.

REFERENCE COUNT: 46 THERE ARE 46 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L2 ANSWER 6 OF 15 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 2002:352833 CAPLUS

DOCUMENT NUMBER: 137:90855

TITLE: Activation and evasion of the antiviral 2'-5' oligoadenylate synthetase/ribonuclease L pathway by hepatitis C virus mRNA

AUTHOR(S): Han, Jian-Qiu; Barton, David J.

CORPORATE SOURCE: Department of Microbiology, University of Colorado Health Sciences Center, Denver, CO, 80262, USA

SOURCE: RNA (2002), 8(4), 512-525
CODEN: RNARFU; ISSN: 1355-8382

PUBLISHER: Cambridge University Press

DOCUMENT TYPE: Journal

LANGUAGE: English

ABSTRACT:

Chronic hepatitis C virus (HCV) infections are a significant cause of morbidity and mortality worldwide. Interferon- α 2b treatment, alone or in combination with ribavirin, eliminates HCV from some patients, but patients infected with HCV genotype 1 viruses are cured less frequently than patients infected with HCV genotype 2 or 3 viruses. The authors report that HCV mRNA was detected and destroyed by the interferon-regulated antiviral 2'-5' oligoadenylate synthetase/RNase L pathway present in cytoplasmic exts. of HeLa cells. RNase L cleaved HCV mRNA into fragments 200 to 500 bases in length. RNase L cleaved HCV mRNA predominately at UA and UU dinucleotides within loops of predicted stem-loop structures. HCV mRNAs from relatively interferon-resistant genotypes (HCV genotypes 1a and 1b) have fewer UA and UU dinucleotides than HCV mRNAs from more interferon-sensitive genotypes (***HCV*** genotypes 2a, 2b, 3a, and 3b). HCV 2a mRNA, with 73 more UA and UU dinucleotides than HCV 1a mRNA, was cleaved by RNase L more readily than HCV 1a mRNA. In patients, HCV 1b mRNAs accumulated silent mutations preferentially at UA and UU dinucleotides during interferon therapy. These results suggest that the sensitivity of HCV infections to interferon therapy may correlate with the efficiency by which RNase L cleaves HCV mRNA.

REFERENCE COUNT: 61 THERE ARE 61 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L2 ANSWER 7 OF 15 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 2001:702106 CAPLUS

DOCUMENT NUMBER: 136:261656

TITLE: Hepatitis C virus (HCV) genotypes, human leukocyte antigen expression and monoclonal gammopathy prevalence during chronic HCV infection

AUTHOR(S): Perrone, Antonio; Deramo, Maria Teresa; Spaccavento, Felice; Santarcangelo, Pasquale; Favoino, Biagio; Antonaci, Salvatore

CORPORATE SOURCE: Department of Internal Medicine, Immunology and Infectious Diseases, University of Bari Medical School, Bari, 70124, Italy

SOURCE: Cytobios (2001), 106(Suppl. 1), 125-134
CODEN: CYTBAI; ISSN: 0011-4529

PUBLISHER: Faculty Press

DOCUMENT TYPE: Journal

LANGUAGE: English

ABSTRACT:

Patients with chronic hepatitis C virus (HCV) infection (530 in toto), and 294 individuals with chronic liver disease of different etiol., were enrolled in this study to investigate the prevalence of monoclonal gammopathies (MG) during

chronic liver dysfunction. A monoclonal band was detected in 61 HCV+ patients and in nine HCV- subjects only. In both instances, a correlation between MG presence and advanced age or degree of hepatic injury was noted. The prevalence of **HCV genotype 2a** was higher in HCV+ patients with, rather than in those without, MG. The MG+ HCV+ subjects did not exhibit human leukocyte antigen (HLA)-A33, B8, B65 and DR16 expression, while an increased frequency of DR15 structure was seen in the same group of individuals in comparison with MG- HCV+ patients and healthy donors. These findings suggest a possible relation between HLA haplotype expression, virus genotypes and the occurrence of MG during the course of chronic HCV infection.

REFERENCE COUNT: 28 THERE ARE 28 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L2 ANSWER 8 OF 15 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 2000:773715 CAPLUS

DOCUMENT NUMBER: 134:325031

TITLE: The relationship between virological characteristics of hepatitis C virus (HCV) and reactivity to the regional specific proteins of HCV

AUTHOR(S): Yoon, Seung Kew; Park, Young Min; Byun, Byung Hoon; Bae, Si Hyun; Yang, Jin Mo; Ahn, Byung Min; Lee, Young Sok; Lee, Chang Don; Sun, Hee Sik; Kim, Boo Sung

CORPORATE SOURCE: Department of Internal Medicine, College of Medicine, Research Institute of Molecular Genetics, WHO Collaborating Center for Reference and Research on Viral Hepatitis, Catholic University of Korea, Seoul, 137-701, S. Korea

SOURCE: Korean Journal of Internal Medicine (2000), 15(2), 109-116

CODEN: KJIMEH; ISSN: 1226-3303

PUBLISHER: Korean Association of Internal Medicine

DOCUMENT TYPE: Journal

LANGUAGE: English

ABSTRACT:

Although the polyproteins of hepatitis C virus (HCV) are processed and formed in nearly equimolar amts., individual functional proteins have a discrepancy in their time of appearance following HCV infection and eliciting immune response. This study was conducted to compare the reactivity toward regional specific HCV protein in relation to virol. characteristics, including HCV genotype and HCV replication. Sera from forty-five patients with chronic HCV infection were analyzed through the expts. of the recombinant immunoblot assay (RIBA-2), HCV genotyping and HCV RNA quantitation. The frequencies of seropositivity to C22-3, C33C, C100-3 and 5-1-1 proteins were 91.1%, 91.1%, 64.4% and 53.3%, resp., of all the patients, and thus the antibodies to C22-3 and C33C proteins were found more frequently ($p < 0.05$). The antibody responses between core or NS3 proteins and NS4 proteins showed more discrepancy in the HCC group than that in the CH group, implying a possibility of oncogenic potential of core or NS3 gene in hepatocarcinogenesis. The detection rate of antibodies to C22-3 and C33C, in accordance with serum HCV RNA levels, was significantly higher in highly viremic patients than that in low viremic patients ($p < 0.05$). Antibodies to C22-3, C33C, C100-3 and 5-1-1 were also found more frequently in patients with HCV genotype 1b, compared to those with **HCV**

genotype **2a** ($p < 0.05$). These results suggest that antibody detection of HCV may depend on the virol. characteristics of HCV, the levels of HCV replication and HCV genotype and, therefore, HCV RNA detection using RT-PCR technique is essential for confirmatory diagnosis for HCV infection. Furthermore, the HCV core or NS3 protein may play important role in hepatocarcinogenesis.

REFERENCE COUNT: 37 THERE ARE 37 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L2 ANSWER 9 OF 15 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 2000:720362 CAPLUS
 DOCUMENT NUMBER: 133:247736
 TITLE: Accurate quantification of HCV-RNA with 2nd generation bDNA probe assay
 AUTHOR(S): Sunaga, Hiroyuki
 CORPORATE SOURCE: Marketing Planning Dep., Diagnostics Div., Daiichi Pure Chemicals Co., Ltd., 13-5, Nihombashi 3-Chome, Chuo-ku, Tokyo, 103-0027, Japan
 SOURCE: Seibutsu Shiryo Bunseki (1999), 22(4), 323-330
 CODEN: SSBUEL; ISSN: 0913-3763
 PUBLISHER: Seibutsu Shiryo Bunseki Kagakkai
 DOCUMENT TYPE: Journal; General Review
 LANGUAGE: Japanese
 ABSTRACT:
 A review with 19 refs. The HCV-RNA quantification method with bDNA probe assay was approved for use under national health insurance in Oct. 1994. In the same time, QuantiplexTM HCV-RNA (version 1.0) was released for sale so that it became possible to use the kit at generic routine test labs. in Japan. Since viral load was recognized as a useful prognosis factor for the Interferon treatment of type C hepatitis, quant. HCV-RNA test with bDNA probe assay has been widely used in the clin. field, and many clin. investigations have been conducted with this method. In the same time improvement of the target specific probe set has been made as the first version was pointed out to underestimate **HCV genotype 2a** and 2b. Now, it has been shown that the improved bDNA probe assay with the new target probe set equally reacts with each HCV genotype transcript. The new bDNA probe test kit (version 2.0) with the new probe set allows us to quantify all of those genotypes within 15 fold range. No matter what the genotype is, accurate quantification of HCV-RNA is now achievable. It is expected that the kit will be a useful tool for the treatment of type C hepatitis with Interferon. This 2nd generation bDNA probe assay kit will be launched in Oct. 1999 in Japan. It is expected that the therapy of type C hepatitis will be based on more accurate viral load quantification.

L2 ANSWER 10 OF 15 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 2000:680723 CAPLUS
 DOCUMENT NUMBER: 134:206368
 TITLE: Predictors of the efficacy of intravenous natural interferon-beta treatment in chronic hepatitis C
 AUTHOR(S): Fukutomi, Takayoshi; Fukutomi, Marie; Iwao, Masataka; Watanabe, Hironobu; Tanabe, Yuichi; Hiroshige, Kaichiro; Kinukawa, Naoko; Nakamuta, Makoto; Nawata, Hajime
 CORPORATE SOURCE: Department of Medicine, Fukuoka City Hospital, Fukuoka, Japan
 SOURCE: Medical Science Monitor (2000), 6(4), 692-698
 CODEN: MSMOFR; ISSN: 1234-1010
 PUBLISHER: Medical Science International Publishing
 DOCUMENT TYPE: Journal
 LANGUAGE: English
 ABSTRACT:
 Several pretreatment factors have been reported to be useful in predicting patients with a high probability for a sustained response to IFN-alpha treatment, however, predictors of the efficacy of interferon-beta treatment in chronic hepatitis C have not been fully assessed. To clarify this issue, a prospective study of 52 patients with chronic hepatitis C was conducted. Patients were treated with human natural interferon-beta by drip infusion at doses of 6 MU/day for 8 wk. The following characteristics were compared between patients with sustained response (SR) and no response (NR): gender, age, source of HCV infection, mean pretreatment serum ALT levels, liver histol., pretreatment serum HCV-RNA levels and HCV genotype. Seventeen of 52 patients (32.7%) demonstrated SR. The proportion of patients with undetectable HCV-RNA levels determined by branched DNA assay (<0.5 + 10⁶ eq/mL) was higher

in patients with SR than in those with NR (88.2% vs. 22.9%; $p=0.0001$). Pretreatment HCV RNA levels determined by multicyclic reverse transcriptase polymerase chain reaction were lower in patients with SR than in those with NR (105.1 ± 1.5 vs. 107.1 ± 1.3 copies/mL; $p=0.0001$). The rate of SR was higher in patients with genotype 2a or 2b than in genotype 1b (43.8% vs. 15.0%; $p=0.0382$). Multivariate stepwise logistic regression anal. showed that a younger age and low pretreatment serum levels of HCV RNA were independent predictors of SR to treatment. This prospective study demonstrated that a younger age, low pretreatment viral load and **HCV genotype ***2a***** or 2b were factors influencing the SR to interferon-beta treatment, but a younger age and low pretreatment viral load were most important predictors of the efficacy of the treatment.

REFERENCE COUNT: 53 THERE ARE 53 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L2 ANSWER 11 OF 15 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 2000:90410 CAPLUS

DOCUMENT NUMBER: 133:100214

TITLE: Detection and molecular cloning of the extreme 3'-end of HCV

AUTHOR(S): Tanaka, Torahiko; Shimotohno, Kunitada

CORPORATE SOURCE: Virology Division, National Cancer Center Research Institute, Tokyo, Japan

SOURCE: Methods in Molecular Medicine (1998), 19(Hepatitis C Protocols), 373-380

CODEN: MMMEFN

PUBLISHER: Humana Press Inc.

DOCUMENT TYPE: Journal

LANGUAGE: English

ABSTRACT:

The 3'-end region of the hepatitis C virus (HCV) genome is composed of three characteristic elements: conventional 3'-untranslated region (3'UTR), subsequent poly U stretch, and a newly identified 3'-terminal sequence named the 3'X tail. The 3'X tail sequence is roughly divided into two categories for convenience, type A including HCV genotypes 1a and 1b and others; and type B including **HCV genotypes 2a, 2b** and others. Here

we present two sets of primers, specific for types A and B, used for amplification of the HCV 3'X tail. We also describe the methods used to clone the 3'X tail of HCV, which include primer extension and RNA ligation expts.

REFERENCE COUNT: 5 THERE ARE 5 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L2 ANSWER 12 OF 15 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 1999:160031 CAPLUS

DOCUMENT NUMBER: 130:336750

TITLE: Interferon therapy for patients with liver cirrhosis due to hepatitis C virus

AUTHOR(S): Miyake, Kazuhiko; Takikawa, Hajime; Yamanaka, Masami

CORPORATE SOURCE: Department of Medicine, Teikyo University School of Medicine, Tokyo, 173-0003, Japan

SOURCE: International Congress Series (1998), 1163(Progress in Hepatology, Volume 4: Liver Cirrhosis Update), 125-132

CODEN: EXMDA4; ISSN: 0531-5131

PUBLISHER: Elsevier Science B.V.

DOCUMENT TYPE: Journal

LANGUAGE: English

ABSTRACT:

We treated 31 patients (33-57 yr old) with liver cirrhosis due to HCV by IFN α . The total IFN dose and duration was 9,878 (175-2,403) MU and 14.8 (3-62) months. The response was defined as complete in cases with normal ALT and neg. HCV-RNA at the end and during the post-treatment period of 6 mo or more. The rate of a complete response was 20 and 33% in HCV genotype 1b and

HCV **genotypes 2a and 2b, resp., and the response**
rate in patients with low HCV-RNA titers was significantly higher than that with high HCV-RNA titers. A complete response required both total doses of more than 500 MU and therapy duration of more than 15 mo. During the average of 44 mo after the IFN therapy, hepatocellular carcinomas were detected in 14% of complete responders and in 25% of nonresponders. According to these findings, an IFN treatment regimen recommended to HCV-related cirrhotics is considered to be based on long-term treatment (at least more than 12 mo), an intermittent therapy (3 times a week), and a proper IFN dose which can maintain neg. serum HCV-RNA for at least 1 yr.

REFERENCE COUNT: 12 THERE ARE 12 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L2 ANSWER 13 OF 15 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 1998:212281 CAPLUS

DOCUMENT NUMBER: 128:304481

TITLE: Usefulness of hepatitis C virus RNA counts by second generation HCV bdna-probe in chronic hepatitis C based on the HCV genotype

AUTHOR(S): Kobayashi, Mariko; Kumada, Hiromitsu; Arase, Yasuji; Chayama, Kazuaki; Kobayashi, Masahiro; Tsubota, Akihito; Koida, Isao; Saitoh, Satoshi; Suzuki, Yosiyuki; Murashima, Naoya; Ikeda, Kenji; Miyano, Yukiko; Mizoshita, Kaori; Matsuda, Marie; Koike, Hiromi; Hashimoto, Michie

CORPORATE SOURCE: Liver Research Laboratory, Toranomon Hospital, Kawasaki, Kanagawa, 213, Japan

SOURCE: Journal of Gastroenterology (1998), 33(2), 223-228
CODEN: JOGAET; ISSN: 0944-1174

PUBLISHER: Springer-Verlag Tokyo

DOCUMENT TYPE: Journal

LANGUAGE: English

ABSTRACT:

Detection of hepatitis C virus (HCV) RNA by a second generations (ver 2) HCV bdna-probe method (bdna-probe) was compared with detection by the first generation (ver 1) assay. The two assays were performed simultaneously with the same serum samples of HCV genotypes 1b, 2a, 2b, 3a, and 3b. The pos. rates with ver 1 were 82% for HCV genotype 1b (type 1b), 57.6% for HCV ***genotype*** 2a (type 2a), 75.0% for HCV genotype 2b (type 2b), 55.6% for HCV genotype 3a (type 3a), and 93.8% for HCV genotype 3b (type 3b). The pos. rates with ver 2 were 95.0% for type 1b, 93.9% for type 2a, 83.3% for type 2b, 100% for type 3a, and 93.8% for type 3b. With Fisher's exact test, the detection rate for type 2a was significantly higher ($P = 0.001$) with ver 2 than with ver 1. We obtained regression lines using the HCV counts measured by bdna-probe on the y axis and the HCV counts obtained by an HCV reverse transcriptase (RT)-competitive polymerase chain reaction method (competitive PCR) on the x axis. The gradients for types 1b, 2a, and 3b were greater with ver 2 compared to ver 1. The gradients for types 2a and 3b were the highest: for type 2a, $y = 0.135x + 0.6$ with ver 1 and $y = 0.248x + 0.1$ with ver 2; for type 3b, $y = 0.366x + 0.1$ with ver 1 and $y = 0.727x + 0.3$ for ver 2. In addition, HCV-RNA counts for all the genotypes tested in this study were significantly higher with ver 2 than with ver 1. Hence, we conclude that ver 2 of the bdna-probe measures HCV-RNA counts closer to those obtained with competitive PCR than the ver 1 assay.

REFERENCE COUNT: 15 THERE ARE 15 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L2 ANSWER 14 OF 15 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 1997:295880 CAPLUS

TITLE: Hepatitis C virus infection in mixed cryoglobulinemia and B-cell non-Hodgkin's lymphoma: evidence for a pathogenetic role

AUTHOR(S): Zignego, A. L.; Ferri, C.; Giannini, C.; La Civita, L.; Careccia, G.; Longombardo, G.; Bellesi, G.; Caracciolo, F.; Thiers, V.; Gentilini, P.

CORPORATE SOURCE: Istituto di Medicina Interna and Hematology Unit, University of Florence, Florence, Italy

SOURCE: Archives of Virology (1997), 142(3), 545-555
CODEN: ARVIDF; ISSN: 0304-8608

PUBLISHER: Springer

DOCUMENT TYPE: Journal

LANGUAGE: English

ABSTRACT:
We investigated the pathogenetic relevance of hepatitis C virus (HCV) infection in mixed cryoglobulinemia (MC) with or without complicating B-cell Non-Hodgkin's lymphoma (NHL) in comparison with other immunol. and lymphoproliferative disorders. The following groups of patients were studied: A) 25 patients with MC in 7 cases evolved into B-cell NHL; B) 25 healthy subjects; C) 22 patients with different systemic immune diseases; D) 24 patients with chronic HCV infection without MC; E) 25 patients with B-cell idiopathic NHL. Methods used included: i) Polymerase chain reaction (PCR) for HCV RNA detection in serum and peripheral blood mononuclear cells (PBMC) (uncultured or mitogen-stimulated); ii) Branched DNA (b-DNA) for HCV RNA quantification; iii) HCV genotyping by genotype-specific primers localized in the core region and by hybridization of amplification products of the 5' untranslated region (5'UTR), obtained with universal primers, using genotype-specific probes. Serum anti-HCV and HCV RNA were detected in 88% and 73% of MC patients, resp., and in a significantly lower percentage of healthy controls and patients with autoimmune diseases. HCV RNA concentration was significantly lower in supernatants than in corresponding whole sera ($p < 0.001$). Plus-strand HCV RNA was detected in 81% of peripheral blood mononuclear cell (PBMC) samples and minus-strand in the majority of fresh or mitogen stimulated cells. All MC patients with NHL had HCV RNA sequences in PBMC. **HCV genotype 2a/III** was detected in MC patients with a prevalence that was significantly higher than in HCV infected patients without MC. Surprisingly, HCV markers (anti-HCV and/ or HCV RNA) were found in 32% of patients with idiopathic NHL. These data suggest that HCV infection is involved in the pathogenesis of MC through both direct participation in the immune complex related vasculitis and by triggering the lymphoproliferative disorder underlying the disease. This latter disorder seems to be related to HCV lymphotropism which could also be responsible for the evolution of MC to malignant lymphoma. This study also suggests that HCV infection may be involved in the pathogenesis of idiopathic B-cell NHL through a similar pathogenetic mechanism.

L2 ANSWER 15 OF 15 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 1995:872870 CAPLUS

DOCUMENT NUMBER: 124:6753

TITLE: Randomized trial comparing three different regimens of alpha-2a-interferon in chronic hepatitis C

AUTHOR(S): Chemello, Liliana; Bonetti, Paola; Cavalletto, Luisa; Talato, Fabiola; Donadon, Valter; Casarin, Pietro; Belussi, Fabio; Frezza, Mario; Noventa, Franco; et al.

CORPORATE SOURCE: Clinica Medica 2, University Padova, Padua, 35100, Italy

SOURCE: Hepatology (Philadelphia) (1995), 22(3), 700-6
CODEN: HPTLTD9; ISSN: 0270-9139

PUBLISHER: Saunders

DOCUMENT TYPE: Journal

LANGUAGE: English

ABSTRACT:
Alpha-interferon (IFN- α) is an effective treatment for chronic hepatitis C, but only 20% to 30% of patients are apparently cured with the currently recommended schedule of 3 MU given three times a week for 6 mo. To evaluate the efficacy of more aggressive treatment regimens, we have conducted a

randomized trial in 174 patients with chronic hepatitis C using three different schedules:. (1) 12-Mo treatment starting with 6 MU/three times a week and decreasing the dose on the basis of serum alanine transaminase (ALT) activities (group A: 59 cases);. Patients were evaluated during therapy for biochem. and virol. response and followed for at least 12 mo after therapy to assess long-term efficacy and liver histol. outcome. The genotype of infecting HCV was also analyzed in all patients, and predictors of response were determined by multivariate anal. Serum ALT became normal during therapy in 76% of patients (95% confidence interval [CI]: 63 to 86), 65% (CI: 52 to 77), and 74% (CI: 60 to 85) in groups A, B, and C, resp. (P = NS). The corresponding figures for sustained response 12 mo after therapy were 49% (CI: 36 to 62), 31% (CI: 20 to 44), and 28% (CI: 16 to 42) (A vs. B, P = .06; A vs. C, P = 0.03). Eighty-six percent of patients with sustained response cleared HCV-RNA from serum, and 72% improved histol. Patients infected with **HCV genotypes ***2a***** and 3 had higher sustained response rates than those with 1b, independent of treatment schedule. In patients infected with genotype 1b, the rate of sustained response was related to dose and duration of therapy being 28% with schedule A, 16% with schedule B, and 9% with schedule C. Multivariate anal. indicated that younger age (P = .016), shorter disease duration (P = .003), and infection with **HCV genotypes 2a** (P = .0017) and 3 (P = .0083) were independent predictors of sustained response. These results indicate that sustained response to IFN- α in chronic hepatitis C is affected by dose and duration of therapy, particularly in patients infected with HCV genotype 1b



Entrez	PubMed	Nucleotide	Protein	Genome	Structure	OMIM	PMC	Journals	E
Search PubMed		for					Preview	Go	
		Limits		Preview/Index	History	Clipboard	Details		

- Search History will be lost after eight hours of inactivity.
- To combine searches use # before search number, e.g., #2 AND #6.
- Search numbers may not be continuous; all searches are represented.

Entrez
PubMed

Search	Most Recent Queries	Time	Result
#14	Search HCV genotype 2a isolate Limits: Publication Date to 2000/06/02	14:16:40	<u>17</u>
#13	Search HCV genotype 2a and cell culture Limits: Publication Date to 2000/06/02	14:16:28	<u>0</u>
#12	Search HCV genotype 2a and cell line Limits: Publication Date to 2000/06/02	14:16:07	<u>3</u>
#11	Search HCV genotype 2a Limits: Publication Date to 2000/06/02	14:15:37	<u>319</u>
#9	Search Bartenschlager R 1999 Field: All Fields , Limits: Publication Date to 2000/06/02	14:12:06	<u>5</u>
#8	Search Bartenschlager R 1999	14:11:02	<u>8</u>
#7	Search Bartenschlager R 1996 Limits: Publication Date to 1999/06/11	14:10:25	<u>4</u>
#6	Search Limits: Publication Date to 1999/06/11	14:09:57	<u>0</u>
#4	Search Lohmann V Limits: Publication Date to 1999/06/11	14:07:42	<u>10</u>
#3	Search Lohman V Field: All Fields , Limits: Publication Date to 1999/06/11	14:07:16	<u>29</u>
#2	Search Lohman V	14:06:19	<u>34</u>
#1	Search Lohman V and HCV	14:06:10	<u>0</u>

PubMed
Services

Related
Resources

Clear History

[Write to the Help Desk](#)
[NCBI](#) | [NLM](#) | [NIH](#)
[Department of Health & Human Services](#)
[Freedom of Information Act](#) | [Disclaimer](#)

JUN 20 2004 11:06:34

0

CORRESPONDENCE ADDRESS:
 ADDRESSEE: HOWELL & HAFERKAMP, L.C.
 STREET: 7733 FORSYTH BLVD., SUITE 1400
 CITY: ST. LOUIS
 STATE: MO
 COUNTRY: USA
 ZIP: 63105
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/034,756
 FILING DATE: 04-May-1998
 CLASSIFICATION: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: HOLLAND, DONALD R.
 REGISTRATION NUMBER: 35,197
 REFERENCE/DOCKET NUMBER: 6029-4831
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 314-727-5188
 TELEFAX: 314-727-6092
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 12980 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cdna
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 SEQUENCE DESCRIPTION: SEQ ID NO: 5:

US-09-034-756-5

Query Match 47.6%; Score 4620; DB 4; Length 12980;
 Best Local Similarity 67.8%; Pred. No. 0;
 Matches 6593; Conservative 0; Mismatches 3035; Indels 92; Gaps 6;

QY	6	CCCTAATAGGGGGACACTCCCGCATGATCACTCCCTGTGAGGAACTACTGTCTTCA	65
DB	7	CCCTGATGGGGGACACTCCACCATGAATCACTCCCTGTGAGGAACTACTGTCTTCA	66
QY	66	CGCAGAAAGCGTCTAGCCATGGCGTGTAGTATGATGTCGTACAGCTCCAGGCCCCCCCC	125
DB	67	CGCAGAAAGCGTCTAGCCATGGCGTGTAGTATGATGTCGTACAGCTCCAGGCCCCCCCC	126
QY	126	TCCCGGAGAGCCATAGTGTCTGCGGAAACCGGTGAGTACACCGGAAATTGCCGGGAAGAC	185
DB	127	TCCCGGAGAGCCATAGTGTCTGCGGAAACCGGTGAGTACACCGGAAATTGCCGGAGCAGC	186
QY	186	TGGGTCTTCTTGGATAAACCCACTCTATGCCCGGCCATTTGGGCGTCCCGCGCAAGA	245
DB	187	CGGGTCTTCTTGGATAAACCCCGCTCAATGCGCTGGAGATTGGGGGTGCGCCCGCAAGA	246
QY	246	CTGCTAGCCAGTAGCTTGGTTGGAAAGGCTTGTGTACTGTCTGATAGGGTCTT	305
DB	247	CTGCTAGCCAGTAGTGTGGTTCGCGAAAGGCTTGTGTACTGTCTGATAGGGTCTT	306
QY	306	GCAGTGTCCCGGGAGTCTCGTAGACCGTGCACCATGAGCACAATCTCTAAACCTCAAA	365
DB	307	GCAGTGTCCCGGGAGTCTCGTAGACCGTGCACCATGAGCACAATCTCTAAACCTCAAA	366
QY	366	GAATAACCAAGAAACCAACACCGTCCGCCACAGACGTTAAGTTCCGGGCGGGCC	425
DB	367	GAATAACCAAGAAACCAACACCGTCCGCCACAGACGTTAAGTTCCGGGCGGGCC	426
QY	426	AGATCGTGTGGAGTATCTTGTTCGCGGAGGGCCCCAGGTTGGGTGTGCGCGCGA	485
DB	427	AGATCGTGTGGAGTATCTTGTTCGCGGAGGGCCCCAGGTTGGGTGTGCGCGCGA	486
QY	486	CAAGGAAGACTTCGGAGCGGTCCCGACGTCGTGGAAAGGCGCCACCTCTCTAAAGATC	545

DB	487	CGAGGAAGACTTTCGAGCGGTGCAACCTCGAGGTAGAGCTCAGCTATATCCCAAGGCAC	546
QY	546	GGCGCTCCACATGGGCAAAATCCTGGGAAACACAGGATACCCCTGGGCCCTATACGGGAATG	605
DB	547	GTCGGCCCGAGGGCAGGACCTGGGCTCAGCCGGGTACCTTTGGGCCCTCTATGGAATG	606
QY	606	AGGGACTCGGCTGGCAGGATGGCTCTGTGTCCCGGAGGTTCCCGTCCCTCTTGGGGCC	665
DB	607	AGGGTTGGGGTGGCGGATGGCTCTGTCTCTCCCGTGGCTCTCGGCTAGCTGGGGCC	666
QY	666	CCAAATGACCCCGGCATAGTTCGCGCAAGTGGGTAAAGTCAATCATACCTTAAAGTGG	725
DB	667	CCACAGACCCCGGCATAGTTCGCGCAATTTGGGTAAAGTCAATCATACCTTAAAGTGG	726
QY	726	GCTTTGCGCACTCATGGGTACATCCCTGTCTGGGGCGCCCGCTCGCGCGCTGCACA	785
DB	727	GCTTCCCGCACTCATGGGTACATACCGTCTGGGGCGCCCTCTTGGAGGCGCTGCACA	786
QY	786	GAGCTCTCCGCAATGGCGTGAAGTCTCTGGAGGACGGGTTAAATTTTCAACAGGAACT	845
DB	787	GGGCGCTGGCGCATGGGTCGGGTTCTGGAAGACGGGTGAATATGCAACAGGGAAC	846
QY	846	TACCGGTTGCTCTTTTCTATCTTCTGTGGGCCCTGTCTCTGCTGCAATCACCACCCGG	905
DB	847	TTCTCTGGTTGCTCTTTCTCTATCTTCTCTTCTGGGCCCTGTCTCTTGGCTGACCGCCG	906
QY	906	TCTCCGCTGCGAAGTGAAGAACATCAGTACCGGCTACATGGTGAATCAACGACTGCACA	965
DB	907	CTTCAGCTTACCAAGTGGCAATTCCTCGGGCTTTTACCATGTCAACATGATGCGCTA	966
QY	966	ATGACAGCAATTACCTGGCAGCTCCAGGCTGTCTCTCCACGTCCTCCGGGTGCTCCCT	1025
DB	967	ACTCGAGTATTGTGTACGAGGCGCGGATGCCATCTCTGCACACTCCGGGTGTGCTCTT	1026
QY	1026	GGGAAAGTGGGAAATGCACTCAGTGTGTGATACCGGTCTCACCAGATGTGGCGTGC	1085
DB	1027	GGTTCGAGAGGTAAACGCTCGAGGTGTGGGTGGCGGTGACCCCGGCGGACCA	1086
QY	1086	ACGCGCCCGGCGCTCAGCAGGCGCTTGCAGACGACATCGACATGGTTGTGATGTCG	1145
DB	1087	GGGAGCGCAACTCCCAACAGCAGCTTCGAGCTCATATCATCTGTTGTGGGAGG	1146
QY	1146	CCAGCTCTGTCTGCTCTAGCTGGGGACCTCTGCGGTGGGTGATGTCTCGAGGCC	1205
DB	1147	CCACCTCTGTCTGCTCTAGCTGGGGACCTCTGCGGTGCTGTCTTCTTGTGGTCT	1206
QY	1206	AAATGTTCAATGTCTCGCGGACGACACTGTGTTCGAGCTGCAAGTGAATGCTCATCT	1265
DB	1207	AACTGTTTACCTTCTCTCCAGGCGCCACTGGACGACGCAAGACTGCAATTTGTTCTATCT	1266
QY	1266	ACCTGGTACCATCACTGACACCGCATGGCATGGGACATGATGAATGCTGCGCCA	1325
DB	1267	ATCCCGGCATATAACGGGTCAATCGCATGCGATGGGATATGATGATGAATGCTCCCTA	1326
QY	1326	CGGCTACCATGATCTTTGGCGTACGCGATGCTGCTCCCGAGGTCAATTATAGACATCTTA	1385
DB	1327	CGGAGCGTGTGTGTAGCTCAGCTGCTCCGATCCCAACAGCCATCATGACATGATCG	1386
QY	1386	GGGGGCTCATTTGGGGCGTCAATGTTGGTGGCCCTACTTCTCTATGAGGAGGCGTGG	1445
DB	1387	CTGTGCTCATCTGGGGAGTCTTGGCGGCGCATAGCGTATTTCTCCATGTTGGGGAATGG	1446
QY	1446	CCAAAGTCTGTCTCATCTTCTTGGCGCGGGGTGAGCGCGCACCCATCTGTTG	1505
DB	1447	CGAAGTCTGTGTAGTGTCTGCTATTTTGGCGGTGCGCGGGAACCCACGTCACCG	1506
QY	1506	GGGGTCTGCGCGCAGACCAACCGGCGCTCACAGCTTATTTGACATGGGCCCCAGGC	1565
DB	1507	GGGGAAGTGGCGCGCAGCCAGCGTGGCTTGTGGTCTCTCTTACACCGCGCCAGC	1566
QY	1566	AGAAATCCAGTCTGTTAAACCAATGAGCAGTGGCACAATCAACCGCACCGGCTGAACT	1625
DB	1567	AGAAATCCAACTGATCAACCAACCGCAGTGGGACATCAATAGCAGCGGCTTGAAT	1626

Copied from 09540843 on 05/19/2004

Db	3775	GAGGTGATAGAGGGTAGCTGCTTTGCGCCCGGCCCAATTTCTACTTTGAAAGGCTCCT	3834
Qy	3846	CAGGAGCCCGGTGCTATGCCCCAGGGGCCACCGTGTGCGAGTCTTCCGGGCGAGCTGTGT	3905
Db	3835	CGGGGGTCCGCTGTGTGCCCCCGGGGACACCGCGTGGGCCCTATTACAGGGCCGCGGTGT	3894
Qy	3906	GCTCTCGGGGCGTGGCTTAAGTCCATAGATTTCAATCCCGTGTGAGACACTCGACATCGTCA	3965
Db	3895	GCACCCGTGGAGTGGCTTAAGCGGTGGACTTTATCCCTGTGGAGAACCTTAGACACCA	3954
Qy	3966	CGCGGTCCCCCACTTTAGTGAACAACAGACACACACTGTGTGTGCCCCAGACCTATCAGG	4025
Db	3955	TGAGATCCCCGGTGTTCACGACAACTCTCTCCACAGCAGTGCCTCCAGAGCTTCCAGG	4014
Qy	4026	TCCGGTACTTGCATGCCCCGACCTGGCAGTGGAGAGGACCAAGTTCCTGTGCGATATG	4085
Db	4015	TGGCCACCTGTCATGCTCCACCGGACGCGTAGAGACCAAGGTCCCGGCTGCGTAGC	4074
Qy	4086	CTGCTCAGGGGTATAAGTGTCTAGTGTCTTAATCCCTCAGTGGCTGCCACCCCTGGGGTTTG	4145
Db	4075	CAGCCAGGCTACAAGTGTGTGCTCAACCCCTCTGTGTGCTGCTGCTGCTGCTGCTGCTG	4134
Qy	4146	GGCGTACTTGTCTAAGGCACTGGCATCAATCCCAACATTAGGACTGGAGTCAGGACTG	4205
Db	4135	GTGCTTACATGTCCAAGGCCCATGGGGTTGATCTTAATATCAGGACCGGGGTGAGAACAA	4194
Qy	4206	TGACGACCGGGCGCCATCAGCTACTCCACATATGCGAAATTCCTCGCGATGGGGCT	4265
Db	4195	TTACCACTGGCAGCCCATCAGGTACTCCACCTACCGCAAGTTCCTGTGCGGCGGGGT	4254
Qy	4266	GTGCGGGCGGCTACGACATCATATGTGATGAATGCCATGCGGTGGACTCTACCA	4325
Db	4255	GCTCAGGAGTGTATGACATTAATTAATGTGACGAGTGCCTCAACGGATGCCACAT	4314
Qy	4326	CCATCTTGTCATCGGAACAGTCTTGATCAAGCAGACAGCTGGGGTCAAGTAACTG	4385
Db	4315	CCATCTTGGGCATCGGCACCTGCTTGACCAAGCAGACAGTGGCGGGCGGAGACTGGTTG	4374
Qy	4386	TGCTGCTCAGCTACGCCCCCTGGGTGAGTGAACACCCCGCCCAACATAGAGAGG	4445
Db	4375	TGCTCGCACTGTGATCCCTCCGGGCTCGCTCACTGTGTCCTCAATCAATCGAGGAG	4434
Qy	4446	TGCGCCCTTGGGAGGAGGCGAGATCCCTTCTATGAGGGGGGATTCCTCCTCTTACA	4505
Db	4435	TTGCTCTGTCCACCAAGGAGATCCCTTTTACGGCAAGGTATCCCTCGAGGTGA	4494
Qy	4506	TCAAGGAGGAGACATCTGATCTTGTCCATTTCAAGAAAGTGTGAGCTCGCG	4565
Db	4495	TCAAGGGGAGACATCTCATCTTCTGCACTCAAGAAAGTGTGCGAGCTCGCG	4554
Qy	4566	CGGCCCTTGGGGTATGGGCTTGAATCTAGTGGCATACTACAGAGGTTGGAGCTCCG	4625
Db	4555	CGAAGCTGTGCTGATGGGCATCAATGCGGTGGCTTACTACCGCGTCTGTGAGTGTCTG	4614
Qy	4626	TAATACCAACTCAGGAGAGCTAGTGTGCTGCGCCACCGCCCTCATGACAGGGTATA	4685
Db	4615	TCATCCCGACCGGGGATGTGTGCTGTGCGACCGATGTCTCATGACTGGGCTTTA	4674
Qy	4686	CTGGGACTTTGACTCGGTGATCGACTGCAACGTAGCGGTCACTCAAGTTGTAGACTTCA	4745
Db	4675	CCGGGACTTCGACTCTGTGATAGATGCAACACGTGTGTCACTCAGACAGTCAATTTCA	4734
Qy	4746	GTTTAGACCCCACTTACCATTAACACACAGATTTGCTTCAAGCGCTGTCTACGTA	4805
Db	4735	GCCTTGACCTTACCTTTACATTTAGACACACCGCTCCCGGAGTGTGTCTCCAGGA	4794
Qy	4806	GCCAGCGCCGGGTGCGCAGGGTAGGGAGAGCTGGGCATTTATAGTATGTTTCCACTG	4865
Db	4795	CTCAAAGCCGGGAGGAGCTGGCAGGGGAGGACGACATCTACAGATTTGTGGCACCG	4854
Qy	4866	GTGAGCGAGCTCAGGAATGTTTGACAGTGTAGTGTCTGTGAGTGTCTACGACGAGGG	4925
Db			
Db	4855	GGGAGCGCCCTCCGCGACTGTTTCGACTCGTCCGTCCTCTGTGAGTGTCTATGACGCGGCT	4914
Qy	4926	CGCATGTGTATGAGCTCACCATCGGAGACCAACCGTCAAGGCTCAGGGCGTATTTCAACA	4985
Db	4915	GTGCTTGGTATGAGCTCAGCCCGCGGAGACTACAGTTAGGCTACAGGCTACATGAACA	4974
Qy	4986	CGCCCGGTTTGCCTGTGTGCCAAGACCACTTTGAGTTTTGGGAGGCGAGTTTTTCCCGGCC	5045
Db	4975	CCCCGGGCTTCCGCTGTGCCAGGACCACTTTGAATTTTGGGAGGCGCTTTTACGGGCC	5034
Qy	5046	TCACACATATAGATGCCCATCTTCTTTCCAAACAAAGCAATCGGGGGGAAATTTTCGAT	5105
Db	5035	TCACATATATAGATGCCCATCTTCTATCCACAAAGCAGAGTGGGGAGAACTTTTCTCT	5094
Qy	5106	ACTTAACAGCTTACAGGCTCAGGTGTGCTAGGSCCAAGGCCCCCCCGCTCTCTGGG	5165
Db	5095	ACTGTAGCTTACAGCCACCGTGTGCGCTAGGGCTCAAGCCCTTCCCCCATCTGGGG	5154
Qy	5166	ACGTATGTGGAAGTGTTTGACTCGACTCAAGCCCACTCGTGGGCCCCCACACTCTCC	5225
Db	5155	ACCAGATGTGGAAGTGTTTGATCCGCTTAAACCCACCCCTCCATGGGCCAACACCCCTGC	5214
Qy	5226	TGTACCGCTTGGGCTCTGTACAGAGGTCACCTTCAACATCCCGTGAAGAAATACA	5285
Db	5215	TATACAGACTGGGCGCTGTTCAGAAATGAAGTCACTGACGCAACCAATCACAAATACA	5274
Qy	5286	TCGCGACCTGACATGCAAGCGGACCTTGAAGTCAAGCAGACATGGGCTTTGGCAGGG	5345
Db	5275	TCATGACATGCAATGTGGCGGACCTTGAAGTCTGACAGCACCTGGGTGCTCTGTTGGG	5334
Qy	5346	GAGTCTTGGCGGCTCGCGGCTATGCTCGCGCACCGGCTGTGTGTGCATCATCGGCC	5405
Db	5335	GGGTCTTGGCTCTCGCGGCTATGCTGTCAACAGGCTCGTGGTTCATAGTGGCA	5394
Qy	5406	GCTTGACATTAACAGGAGGCGCTGTGCGCGCGGACAGAGGCTCTCTATGAGGCTT	5465
Db	5395	GGATGTCTTGTTCGGGAGCGCGCAATTAATACCTGACAGGAGGTTCTTACACAGGAT	5454
Qy	5466	TTGATGAGTGGAGGAATGTGCTCTAGGGCGGCTCTCATTTGAAGGGGCGAGCGATAG	5525
Db	5455	TCGATGAGTGAAGAGTGTCTCAGCACTTACCGTACATCGAGCAAGGATGATGCTCG	5514
Qy	5526	CCGAGATGTGAAGTCCAAGATCCAAGGCTTATGTCAGCAAGCTTCCAAACAAAGCTCAAG	5585
Db	5515	CTGAGCAGTTCAGCAGAGGCGCTCGGCTCTCTGACAGCCGCTCCCGCAGCAGAGG	5574
Qy	5586	ACATCAACCCCATGTGAGGCTTCAATGCGCCCAAGGTAGAACAAATTCGGGGCAACACA	5645
Db	5575	TTATCACCCCTGCTGTCCAGACCAACTGGCAGAAACTCGAGGTCTTCTGGGCGAAGACA	5634
Qy	5646	TGTGGAATTCATTAAGCGGCATCCCAATACCTCGCAGGACTATCAACACTGCCAGGNAAC	5705
Db	5635	TGTGGAATTTTCATAGTGGGATACAATATCTGGCGGGCTGTCAACGCTGCTGTGAACC	5694
Qy	5706	CTGCACTAGCTTTCATGATGCGGTTCACTGCGCGCTTCAACAGTCCGCTGTCAACAGCA	5765
Db	5695	CGCCCATGCTTTCATGATGCTTTTACAGCTGCGCTCACCAGCCCACTAACCACTGGCC	5754
Qy	5766	CCATATCTTCTCAACATTTTGGGGGCTGTGCTAGCTATCCAAATTCGACCAACCGCGG	5825
Db	5755	AAACCTCTCTTCAACATATTTGGGGGGTGGGCTGCCAGCTCGCGCGCCCGGCTG	5814
Qy	5826	GGGCGACTGCTTCTGTGAGTGGGCTAGTGGGAGCTGCGTAGGAGTATAGCTTAG	5885
Db	5815	CGCTACCGCTTGTGGGCGCTGTGCTGTAGTGGCGGCCCATCGGACGCTTGGACTGG	5874
Qy	5886	GTAAAGTGTCTAGTGACATCTCTGGCAGGCTATGTTGGGGGCTTTCGGGGGCTCTCGTCG	5945
Db	5875	GGAGGTCTCTCTGACATCTTGTGAGGATGAGCGCGGGCTGCGGGGAGCTCTGTAG	5934
Qy	5946	CATTCAAGATCATGTCTCGGAGAGGCTTCCATCGAGGATGTCTCAACTTGTCTGCTG	6005
Db	5935	CCTTCAAGATCATGAGCGGTGAGGTCCCTTCCACGAGGAGCCTGGTCAATCTGCTGCCG	5994

QY	6006	GAATTCCTGCTCGGGTGCCTTGGTAGTGGAGTCACTGCGGGCCATCTCTGCGCGGAC	6065
Db	5995	CCATCTCTCGCTGAGCCCTTGTAGTCGGTGTGCGGAGCAATACTGCGCGGC	6054
QY	6066	ACGTGGGACCGGGGAGCGCGCTCCATGGATGAATAGACTCATTTGCTTTGCTTCCA	6125
Db	6055	ACGTTGGCCCGGCGAGGGGAGTGCAATGGATGAACCGGCTAATAGCCTTCGCTCC	6114
QY	6126	GAGGAATCAGCTGCGCCGCCACCACTAGCTGACGGAGTGGATGGTGCAGCGTGTGA	6185
Db	6115	GGGGGAACCATGTTTCCGCCACGACTACGTGCGGAGAGCGATGACGCGCCGCTCA	6174
QY	6186	CCCACTACTTGGCTCCCTTACCATACAGCGCTGCTCAGAAGACTCCCAACTGGATTA	6245
Db	6175	CTGCCACTACAGAGCTCAGTGAACCACTGTAACCACTCTCTGAGCGACTGCATCAGTGATTA	6234
QY	6246	CTGAGGACTGCCCATCGCGGGCTCGTGGCTCCGCGATGTGTGGGACTGGGTTT	6305
Db	6235	GCTCGGAGTGTACCACTCCATGCTCCGGTTCCTGGCTAAGGACATCTGGGACTGGATAT	6294
QY	6306	GCACCATCTTAACAGACTTTAAATTTGGCTGACCTCCAAATTAATTTCCCAAGATGCCG	6365
Db	6295	GCGAGTGTCTGAGCGACTTTAAGACCTGGCTGAAAGCCAAAGCTCATGCCCAACTGCCTG	6354
QY	6366	GCCTCCCTTTGTCTCTGTCAAAAGGGGTACAAGGCGGTGCGCGGCACTGGCATCA	6425
Db	6355	GGATCTCCCTTTGTCTCTGACGGCGGTATAGGGGGTCTGCGAGGAGACGGCATTA	6414
QY	6426	TGACCACACGGTGTCTTGGCGGCCCAATATCTCTGCAATGTCCGCTTGGGCTCCATGA	6485
Db	6415	TGCACACTCGCTGCACTGTGGAGCTGAGATCACTGGACATGTCAAAAACGGACATGA	6474
QY	6486	GAATCACGGGGCTTAAGACCTGCAATGAATATCTGGCAGGGGACCTTCTCTATCAATGTT	6545
Db	6475	GGATCTGCTGCTTAGGACCTGACGGAACATGTGGAGTGGGACGTTCCCAATTAACCGCT	6534
QY	6546	ACAGGAGGCGAGTGTGTCGCGGAAACCCGCGCCAACTTTAAGGTGCGCATCTGGAGG	6605
Db	6535	ACACCGGGCCCTGTACTCCCTTCTGCGCGCAACTAATAGTTCGCGCTGTGGAGG	6594
QY	6606	TGGCGGCTCAGAGTACGCGAGGTGACGAGCACGCGGTCAATCACTAATCAACAGGAC	6665
Db	6595	TGTCGACAGGAATACGTGGAGATAAGGCGGTGGGGGACTTCCACTAGTATCGGTA	6654
QY	6666	TCACCATGATACTTGAAGTCCCTGCGCAACTACCTCTCCCGAGTCTTTTCTCTGGG	6725
Db	6655	TGACTACTGACAACTTTAAATGCGCGTGCAGATCCCATCGCCGAAATTTTTCACAGAAT	6714
QY	6726	TGGACGAGTGCAGATCCATAGGTTTGCCTCCACACCGAGCGCTTTTCCGGGATGAGG	6785
Db	6715	TGGACGGGTGCGCTACATAGTGTGGCGCCCTTGCAGACCTTGTCTCGGGAGGAGG	6774
QY	6786	TCTGTTCTGCGTGGCTTAATTCATTTGTGCTCGGTCCTCAGCTCTCTTGCAGCCCTG	6845
Db	6775	TATCATTCAGATAGGACTCCACAGTACCGGTGGGTGCGCAATTAACCTTGGAGCCG	6834
QY	6846	AACCCGACACAGAGTATGATGTCATGTAAAGATCACTATCATATCACTGCGCGGAGA	6905
Db	6835	AACCGGACGTAGCGCTTGAACGCTCCATGCTCACTGATCTCTCCCATATAACAGCAGG	6894
QY	6906	CTGACGCGCGGTTTACGCGGGGTACCCCATCCGAGGCAAGCTCTCTCGCGAGCC	6965
Db	6895	CGGCGGAGAGAGTGGGAGAGGGTACCCCTTCTATGGCAGCTCTCTCGCCAGCC	6954
QY	6966	AGCTATCGGACCATCTGCGAGCCACCTGACACCCACCGCAAGCCATATGATGTGG	7025
Db	6955	AGCTGTCCGCTCCATCTCTCAAGGCAACTTGCACCGCCCAACCTAGCTCTCTCGCCAGCC	7014
QY	7026	ACATGTTGGATCTAACT-----GTTTCATGGGGGGGATGTGACTCGGATAG	7073
Db	7015	AGCTCATAGGGCTTAACCTCTCTGTGGAGGAGGAGATGGGCGGCAACATCACCAGGGTTC	7074
QY	7074	AGTCTGGGTCCAAAGTGTGCTTGTGGACTCTCTCGAACCAATGTGCGAAGAAAGAGCG	7133
Db	7075	AGTCAGAGAAACAAGTGTGATTTCTGGACTCTCTCGATCCGCTTGTGCGAGGAGGATG	7134
QY	7134	ACCTTGAGCTTCGATACCATCAGATACATGCTCCCCAAGAGAGGTTCACCAACAGCTT	7193
Db	7135	AGCGGAGGTCTCCGTACCCGAGAAATTTCTCGGAAGTCTCGGAGATTCGCGCGGCCCC	7194
QY	7194	TACCGGCTTGGGACGCGCTGATTAACACCCACCGCTTGTGGAATCGTGGAAAGGCCAG	7253
Db	7195	TGCCGCTTTGGCGCGCGCGGACTACACCCCGCTAGTAGAGAGCTGGAAGAAAGCTTG	7254
QY	7254	ATTACCAACCGGCTGCTGTTGCGGCTGTGCTCTCCCTCTCTAGGAAACCCGACGC	7313
Db	7255	ACTACGAACCACTGTGTGCTCATGCTGCGCTACCACTCCACGGTCCCTCTCTGTGC	7314
QY	7314	CTCCCCAAGAGGCGCGGACAGTGGGCTTAAGTGTAGAGCTCCATAGAGAGATGCCCTTC	7373
Db	7315	CTCCGCTTCGGAAGAGCGTACGCTGGTCTCTCACGAATCAACCTCTATCTACTGCTTGG	7374
QY	7374	AACAGTGGCAATTAAGTCTTTTGGCCAGCGCCCCCAAGCGCGATTCAGGCTTTTCCA	7433
Db	7375	CCGAGCTTCCCAACAAAGTTTTTGGCAGCTCTCTCAACTTCCGCGATTACGGGGACAATA	7434
QY	7434	CGGGGCGCGGCTGCGGATTCGCGCAGTCAGACGCTCTCTGATGAGTTTGGGCTTTCG	7493
Db	7435	CGACAACATCTCTG-----AGCGCGCCCTTCTGCTGCCCCCGACTCCG	7482
QY	7494	AGACAGTTCATCTCTTCCATGCCCCCTCGAGGGGAGCTTGGAGATCCAGACTTCG	7553
Db	7483	ACGTTGAGTCTTATCTTCCATGCCCCCTCGAGGGGAGCTTGGGATTCGGATCT--	7540
QY	7554	AGCCTGAGCAGTAGAGCCCCCAACCCCCCGGGGGGTGGCAGCTCCCGGCTCG	7613
Db	7541	-----CAGCGAGC	7548
QY	7614	ACTCGGGTCTGCTTACTTCTCGAGGAGGAGCACTCCGCTGCTGCTGCTCATGT	7673
Db	7549	GGTCATGTCGAGCGGTACGTAGTGGGCGGACACGGAAGATGTGCTGCTGCTCAATGT	7608
QY	7674	CATCTCTGAGACCGGGCTTAATTAATCTTGTAGTCCGAGAGGAGGATTTACCGA	7733
Db	7609	CTTATCTCTGAGCAGCGCACTCGTACCCCGTGGCTCGGAGAAACAAAACCTGCCA	7668
QY	7734	TTAAACCTTTGAGCACTCCCTGTTGCGATATACAAAGGTGTACTGTACCAACAA	7793
Db	7669	TAACGCACTGAGCACTCTGTTGCTAGCCATCACAATCTGTTGTTATTCACCACTTCA	7728
QY	7794	AGAGCGCTCCTCAAGGGCTAAAAAGGTAACTTTTGTATAGGATGCAAGTCTCGACTCT	7853
Db	7729	GCAGTCTTTGCCAAAGGAGAGAAAGTCAATTTGACAGACTGCAAGTTCTGGACAGCC	7788
QY	7854	ACTACGACTGCTTTAAAGGACATTAAGTGTAGCGCTTCCAGGTACCGGAGGCTCC	7913
Db	7789	ATTACGAGGCTGTCTCAAGGAGGTCAAGAGCGCGCTCAAAAGTGAAGGCTTAACTTGC	7848
QY	7914	TCAACATGAGGAGGCTTCCAGTTAAACCCCACTTCTGCAAGATCTTAATATAGGGT	7973
Db	7849	TATCCGTAGGAGGCTTGCAGCTGACGCCCTCCACATTCAGGCAATCCAGTTGGCT	7908
QY	7974	TTGGGGCTAAGGAGTCCGAGCTTGTCCGGAGGGCCGTTTAAACCACTCAAGTCCGCTGT	8033
Db	7909	ATGGGCAAAAGACGTCCTGCTTGCATGCGCAGAAAGCCGTAGGCCACATCAACTCCGCTGT	7968
QY	8034	GGAGGACCTCTGAGGACTCAGAAACCAATTTCCCAACCACTATGGCCAAAATG	8093
Db	7969	GGAAAGACCTTCTGGAAGACAGTGTAAACCAATAGACACTACCATCATGGCCCAAGACG	8028
QY	8094	AGGTGTTCTGCTGGACCCCAACCAAGGGGCGGCAAGAAAGCAGCTCGCTTATCGTTTACC	8153
Db	8029	AGTGTCTGCTGTTAGCTTGAAGAGGGGGTGTAAAGCCAGCTCTCTCATCTGTGTTC	8088
QY	8154	CTGACCTCGCGCTCAGGGTCTGCGAAGAGATGGCCCTTTTATGACATTTACAAAACTTC	8213

	Dd	9159	AGTACTCTTTCAACGTGGGCGAGTAGAACAAGAAGCTCAAATAGCGCGCGTG	9228
	Qy	9294	GCCTCCTGGATTATPCCAGCTGGTTACCGTCGCCGCCGGGGCGGCGCATTTATCAC	9353
	Dd	9229	GCCGGCTGGACTTGTCCGGTTGGTTTCA CGGCTGGCTACAGCGGGGAGACATTTATCAC	9288
	Qy	9354	GGGTGCGGGTCCCCGACCCCCTTAATGCTCTTTTGGCCTACTCCCTACTTTTTGTAGGG	9413
	Dd	9289	GGGTGTCTCATGCCCCGGCCCCGCTGGTTCTGGTTTGGCTACTCTCTGCTCGCTGCAGGG	9348
	Qy	9414	TAGGCTTTTCTCTACTCTCCCGCTCG--GTAGAGCGGCACACATTAAGCTACACTCCATAG	9471
	Dd	9349	TAGGCATCTACTCTCTCCCAACCAGTAAGTTGGGTAAACATCTCCGCCCTCTTAGGC	9408
	Qy	9472	TAACTGTCCCTT	9531
	Dd	9409	CATTTCCCTGTTT	9468
	Qy	9532	TTTCTCTTTCTCTTCTT	9591
	Dd	9469	TTCTTTTTTTTTTTTTTTTTTTCCTTTTTTTTTTTTTTTTTTTTTTTCTTTCTCTTCT	9528
	Qy	9592	ACCTATATTTACTTTCTTTCTCTGGTGGCTCCAATCTTAGCCCTAGTACCGCTAGCTGCA	9651
	Dd	9529	TTCTTTTCTCTCTTTCTTAATGGTGGCTCCAATCTTAGCCCTAGTACCGCTAGCTGCA	9588
	Qy	9652	AAGTCCGTGAGCCGATGACTGCAGAGAGTCCCGTAAC TGTCTCTCTGCAGATCATGT	9711
	Dd	9589	AAGTCCGTGAGCCGATGACTGCAGAGAGTCTGATACTG GCTCTCTGCGAGATCATGT	9648
RESULT 9				
US-09-014-416-6				
; Sequence 6, Application US/09014416				
; Patent No. 6153421				
; GENERAL INFORMATION:				
; APPLICANT: Yanagi, Masayuki				
; APPLICANT: Bukh, Jens				
; APPLICANT: Emerson, Susanne U.				
; APPLICANT: Purcell, Robert H.				
; TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND				
; TITLE OF INVENTION: USES THEREOF				
; FILE REFERENCE: 20264276				
; CURRENT APPLICATION NUMBER: US/09/014,416				
; CURRENT FILING DATE: 1998-01-27				
; EARLIER APPLICATION NUMBER: US 60/053,062				
; EARLIER FILING DATE: 1997-07-18				
; NUMBER OF SEQ ID NOS: 65				
; SOFTWARE: PatentIn Ver. 2.1				
; SEQ ID NO 6				
; LENGTH: 9599				
; TYPE: DNA				
; ORGANISM: Hepatitis C virus				
US-09-014-416-6				
Query Match 46.7%; Score 4534.4; DB 3; Length 9599;				
Best Local Similarity 67.4%; Pred. No. 0;				
Matches 6550; Conservative 0; Mismatches 3031; Indels 137; Gaps 6;				
Qy	6	CCCCTAATAGGGGCGACACTCGCCCATGAATCACTCCCTGTGAGAACTACTGTCTTCA	65	
Dd	7	CCCTGATGGGGGCGACACTCCACCATGAATCACTCCCTGTGAGAACTACTGTCTTCA	66	
Qy	66	GCAGAAAGCGTCTAGCCATGCGGTAGTATGAGTGTGCTAGAGTGTGCTAGGCTCCAGGCCCCCCCC	125	
Dd	67	GCAGAAAGCGTCTAGCCATGCGGTAGTATGAGTGTGCTAGAGTGTGCTAGGCTCCAGGCCCCCCCC	126	
Qy	126	TCCCGGAGAGCCATAGTGGTCTGCGGAAACCGGTGAGTACACCGGAATTGCCGGGAAGAC	185	
Dd	127	TCCCGGAGAGCCATAGTGGTCTGCGGAAACCGGTGAGTACACCGGAATTGCCGAGACGAC	186	
Qy	186	TGGTCTCTTTTGGATAAACCCACTATGCCCCGGCCATTGGGGCGTGCCTCCCGCAAGA	245	

RESULT 9

Copied from 09540843 on 05/19/2004

[illegible]

4555 CGAAGCTGTGCGCATTTGGGCATCAATGCCGTGGCTACTACCGCGTCTTGACGTGCTG 4614
 4626 TAATACCAACTCAGGAGAGGTAGTGTGCTGCTCCACCGACGCGCTCATGACAGGCTATA 4685
 4615 TCATCCCGACACCGGCGATGTTGTGCTGTGTCGACCGATGCTCTCATGACTGGCTTTA 4674
 4686 CTGGGACCTTTGACTCCGATCGACTGCAACCGTAGCGGTCACTCAAGTTGTAGACTTCA 4745
 4675 CCGGCGACTTCGACTCTGTGATAGACTGCAACAGTGTGCTCACTCAGACAGTCGATTTCA 4734
 4746 GTTTAGACCCCACTTCAACCAATCAACCAACAGATGTGCTCAAGCGTGTCTCAGTA 4805
 4735 GCCTTGACCTACCTTTACCAITGAGACCAACACGCTCCCGCAGGATGCTGTCTCCAGGA 4794
 4806 GCCAGCGCGGGTCCGACGGGTAGGGAGACTGGGCATTTATAGGTATGTTCCACTG 4865
 4795 CTCAACCGCGGGCGAGGACTGGCAGGGGGAAGCAGGCATCTATAGATTGTGGCACCGG 4854
 4866 GTGACGAGGCTCAGGAATGTTTGACAGTGTAGTGTCTGTGAGTGTCTAGCAGCGAGGG 4925
 4855 GGGAGCGCCCTCCGCGATGTTGCACTGCTGCTCTCTGTGAGTGTCTATGACGCGGGT 4914
 4926 CCGCATGTTAGTCTACACCATCGGAGACCAACCGTCAAGGCTCAGGGCTATTTCAACA 4985
 4915 GTGCTTGGTATGACTCAGCCCGCGAGACTAGTTAGGCTACGAGCGTACATGAACA 4974
 4986 CGCCCGGTTGCTGTGTCAGACCAATCTTGAGTTTGGGAGGAGTTTTCACCGGC 5045
 4975 CCCCGGGCTTCCGCTGTGCGAGGACCATCTTGAATTTTGGGAGGCGCTTTTACGGGGC 5034
 5046 TCACACATAGATGCCACTCTCTTCCCAACAAAGCAATCGGGGAAATTTTCGCAT 5105
 5035 TCATCATATAGATGCCACTTTTATCCAGACAAAGCAGTGGGAGACTTTTCCTT 5094
 5106 ACTTAAACAGCTACAGGCTACAGTGTGCTAGGCGCAAGCCCCCCCCCTGCTGGG 5165
 5095 ACCTGTAGGCTACCAAGCACCGTGTGCTAGGCGCTCAAGCCCCCTCCCGCATCTGGG 5154
 5166 ACCTGTAGGAGTGTGACTTCGACTCAAGCCACACTCGTGGGCGCCACACTCTCC 5225
 5155 ACCAGATGTGGAAGTGTGATCCGCTTAAACCCACCTCCATGGGCGCAACCCCTGC 5214
 5226 TGTACCGCTTGGGCTGTGTTACCAACAGAGTCAACCTCCACATCCCGTGACCAATACA 5285
 5215 TATACAGACTGGGCGTGTTCAGATGAAGTCACTGACGACCAACCAACCAATACA 5274
 5286 TCGCCACCTGCATGCAAGCGGCTTGTAGTCAATGACAGCAATGGGCTTGGCGAGGG 5345
 5275 TCATGACATGATGTGCGGCGGCTGTGAGGTGCTGACGAGCACCTGGGTCTGTTGGG 5334
 5346 GAGTCTTGGGCGGCTGCGCGGTATTTGCTGGGACCGGGTGTGTTGCATCATCGGC 5405
 5335 CGCTCTGCTGCTGTGGCGCGTATTTGCTGTCAACAGGCTCGGTGATAGTGGGCA 5394
 5406 GCTTTCACATTAACACGAGCGCTGCTGCGCGGACAGAGGCTCTTATGAGGCTT 5465
 5395 GGATGCTTGTTCGGGAACCGGCAATTAATCTGACAGGAGGTTCTCTACAGAGT 5454
 5466 TTGATGAGTGGAGGATGTGCTCTTAGGCGGCTCTCATTTGAAGAGGGGACGCGATAG 5525
 5455 TCGATGAGATGGAAGAGTGTCTCAGCACCTTACCGTACATCGAGCAAGGATGATGCTG 5514
 5526 CCGAGATGCTGAAGTCAAGATCAAGGCTTATGCGAGGCTTCCAAACAGCTCAAG 5585
 5515 CTGAGCAGTTTCAAGCAGAAAGGCTTCCGCGCTCTGCGAGACCGGCTCCCGCATGCAAGG 5574
 5586 ACATACAACCACTGTGCGGCTTCAATGCGCCCAAGGTAGAACAAATTTGGGCGCAACACA 5645
 5575 TTATACCCCTGTGTCAGAACCACTGGCAGAACTCGAGGTTCTTTTGGGCGAGGACA 5634
 5646 TGTGGAATTTTATAGCGGATCCAAATACCTCGAGGACTATCAACACTGCCAGGAAACC 5705
 5635 TGTGGAATTTTATCAGTGGGATCAATATCTTGGCGGCGCTGTCAAAGCTGCTGGTAACC 5694

QY 5706 CTGCACTAGCTTCCATGATGGGCTTCACTGCGCGCTCACCAGTCCGCTGTCAACAAGCA 5765
 Db 5695 CGCCCATGCTTCACTTGTGCTTTTACAGCTGCGCTCACCAGCCCACTAACCACTGGCC 5754
 QY 5766 CCACTATCTTCTCAACATTTTGGGGGCTGCTAGCTAGCTCCAAATTTGACACACCGCGG 5825
 Db 5755 AAAACCTCTCTTCAACATATTTGGGGGCTGCTGCTGCTCCCAAGCTGCGCGCGCGGTG 5814
 QY 5826 GGGCCACTGGCTTCTGCTAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5885
 Db 5815 CGCTACTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5974
 QY 5886 GTAAGTGTCTAGTGGACATCTTGGGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5945
 Db 5875 GGAAGTCTCTGCTGACATCTTTCAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5934
 QY 5946 CATTCAGATCATGTCTGCGAGAGCCCTTCATGAGGATGCTGCTGCTGCTGCTGCTGCTG 6005
 Db 5935 CATTCAGATCATGAGCGGTGAGGTCCCTTCCACGAGGACCTGCTGCTGCTGCTGCTGCTG 5994
 QY 6006 GAAATCTGCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6065
 Db 5995 CCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6054
 QY 6066 AGTGGGACCGGGGAGGCGCGCTCAATGATGAATAGACTCATTTGCTGCTGCTGCTGCTG 6125
 Db 6055 AGTGTGGCGCGGCGAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6114
 QY 6126 GAGGAATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6185
 Db 6115 GGGGGAACCATGCTTTCCTCCCGCAGCTACGCTGCGGAGGAGCTGCTGCTGCTGCTGCTG 6174
 QY 6186 CCAACTACTTGGCTTCCCTTACATTAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6245
 Db 6175 CTGCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6234
 QY 6246 CTGAGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6305
 Db 6235 GCTGCGAGTGTACCTCACTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6294
 QY 6306 GCACCATCTTAAACAGCTTTTAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6365
 Db 6295 GCGAGTGTGAGCGACTTTAAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6354
 QY 6366 GCTTCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6425
 Db 6355 GGAATCCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6414
 QY 6426 TGACACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6485
 Db 6415 TGACACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6474
 QY 6486 GAAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6545
 Db 6475 GGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6534
 QY 6546 ACACGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6605
 Db 6535 ACACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6594
 QY 6606 TGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6665
 Db 6595 TGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6654
 QY 6666 TCACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6725
 Db 6655 TGACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6714
 QY 6726 TGACGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6785
 Db 6715 TGACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6774

8869 TGATAGTATGATGACCCATTTCTTTAGCGTCTCATAGCCAGGGATCAGTTGAACAGGCTC 8928
 8994 TCAACTTTGAGATGATGAGGAGCGGTGACTCCGTTGAGTCCCTTGGACCTCCCGAGCTATAA 9053
 8929 TTAACCTGTGAGATCTACGGAGCGCTGCTACTCCATAGAACCACTGATCTACCTCCATCA 8988
 9054 TTCAAAAGGTTACATGCGCTTACGCTTTTCTCTGACACATACATCTCCCGACGAATGA 9113
 8989 TTCAAAAGATCTCAATGCGCTCAGCGATTTTCACTCCACAGTTACTCTCAGGTGAATCA 9048
 9114 CAGCGGTGGCTTCAGCGCTCAGAAACTTTGGGCGCCACCCCTCAGAGCGTGAAGAGCC 9173
 9049 ATAGGTGGCGCATGCTCAGAAACTTTGGGCTCCGCGCTTGGAGCTTGGAGACACC 9108
 9174 GGCACGTGAGTACAGGCGTCCCTCATCTCCCGTGGGGGAGAGCGCGTTTGGGTC 9233
 9109 GCGCGCGGAGCGTCCGCGCTAGCGTCTGTCAGAGAGGAGCGAGGCTGTATATGGCA 9168
 9234 GATATCTCTTCAATTTGGGCGGTGAAGACCAAGCTCAAACTCACTCCATTTGCCGGAAGCGC 9293
 9169 AGTACCTCTTCACTGGGCACTAAGAACAAAGCTCAAACTCACTCCCAATAGCGCGCTG 9228
 9294 GCTCTCTGATTTATCAGCTGTTTCAACGCTCGCGCGCGCGCGGCGGAGCAATTTATCA 9353
 9229 GCGCGCTGGACTTGTCCGTTGGTTACGGCTGGCTACAGCGGGGAGACATTTATCA 9288
 9354 GCGTGTGCGTGGCGGCGCGCTTATGCTCTTGGCGCTACTCTCTCTTTTGTAGGG 9413
 9289 GCGTGTCTCATGCGCGCGCGCTGCTGTTTGGCTTCTGCTGCTGCTGCGTGCAGGGG 9348
 9414 TAGCGCTTTTCTACTCTCCCGCTCGGTAGAGCGGACACATAGTACACTCCATAGCTA 9473
 9349 TAGCATCTACTCTCTCCCAACGATGAAGTTGGGTAAACACTCCGCGCTCTTAAGC 9408
 9474 ACTGTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9533
 9409 CATTTCTGTTTCC 9468
 9534 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 9593
 9469 TTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 9485
 9594 CTATTTTACTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 9653
 9486 -----TTTCTTTCTTTTAAAGTGGTCTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 9541
 9654 GGTCCGTGAGCGGATGATGAGAGAGTGCATGAGTGCATGAGTGCATGAGTGCATGAGT 9711
 9542 GGTCCGTGAGCGGATGATGAGAGAGTGCATGAGTGCATGAGTGCATGAGTGCATGAGT 9599

RESULT 10
 US-09-014-416-2
 ; Sequence 2, Application US/09014416
 ; Patent No. 6153421
 ; GENERAL INFORMATION:
 ; APPLICANT: Yanagi, Masayuki
 ; APPLICANT: Buhn, Jens
 ; APPLICANT: Emerson, Susanne U.
 ; APPLICANT: Purcell, Robert H.
 ; TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND
 ; TITLE OF INVENTION: USES THEREOF
 ; FILE REFERENCE: 20264276
 ; CURRENT APPLICATION NUMBER: US/09/014,416
 ; EARLIER FILING DATE: 1998-01-27
 ; EARLIER APPLICATION NUMBER: US 60/053,062
 ; EARLIER FILING DATE: 1997-07-18
 ; NUMBER OF SEQ ID NOS: 65
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 9599
 ; TYPE: DNA
 ; ORGANISM: Hepatitis C virus

US-09-014-416-2
 Query Match 46.6%; Score 4523.2; DB 3; Length 9599;
 Best Local Similarity 67.3%; Pred. No. 0;
 Matches 6543; Conservative 0; Mismatches 3038; Indels 137; Gaps 6;
 QY 6 CCCCTAAATAGGCGGACACACTCCGCGCATGAATCACTCCCTCTGTGAGGAACTACTGTCTTCA 65
 DB 7 CCCCTAGTGGGCGGACACTCCACCATGAATCACTCCCTCTGTGAGGAACTACTGTCTTCA 66
 QY 66 CCGAGAAAGCGTCTAGCCATGCGGTTAGTATGAGTGTCTGATGAGTCTCAGCTCCAGGCCCCCCCC 125
 DB 67 CCGAGAAAGCGTCTAGCCATGCGGTTAGTATGAGTGTCTGATGAGTCTCAGCTCCAGGCCCCCCCC 126
 QY 126 TCCCGGAGAGCATAGTGTCTGCGGAACCGGTGAGTACACCGGAATTTCCGGGAAGAC 185
 DB 127 TCCCGGAGAGCATAGTGTCTGCGGAACCGGTGAGTACACCGGAATTTCCCGGAGGAC 186
 QY 186 TGGGTCCTTTTCTTGGATAAACCCACTCTATGCGCGGCCAATTTGGGCGTGCCTCCCGCAAGA 245
 DB 187 CCGGTCCTTTTCTTGGATAAACCCGCTCAATGCTCGAGATTTGGGCGTGCCTCCCGCAAGA 246
 QY 246 CTGCTAGCCGAGTGTGGTTGGGTTGGAAAGCGCTTGTGTTACTGTGCTGATAGGGTGTCTT 305
 DB 247 CTGCTAGCCGAGTGTGGTTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTT 306
 QY 306 GCGAGTCCCGCGGAGTCTGCTAGACCGTGCACCATGAGCACAATCTTAAACCTCAAA 365
 DB 307 GCGAGTCCCGCGGAGTCTGCTAGACCGTGCACCATGAGCACAATCTTAAACCTCAAA 366
 QY 366 GAAACCAAAAGAAACCAACCGTGCCTGCGCCCAAGACCGTTAAGTTTCCGGGCGCGGCC 425
 DB 367 GAAACCAAAAGAAACCAACCGTGCCTGCGCCCAAGACCGTCAAGTTTCCGGGTTGGGTC 426
 QY 426 AGATCTGCGGAGTATATCTTGTGCGCGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 485
 DB 427 AGATCTGCGGAGTATATCTTGTGCGCGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 486
 QY 486 CAGGAGAGATCTCGGAGCGGTCGCGACCGTGGAGGCGCGGCGGCGGCGGCGGCGGCGGCGG 545
 DB 487 CAGGAGAGATCTCGGAGCGGTCGAAACCTCGAGGTAGAGTCAAGCTATCCCCCAAGCAC 546
 QY 546 GCGGCTTCACTGCGCAAACTCTGGGGAACCAAGGATACCCCTGGCGCCCTATACGGGAATG 605
 DB 547 GTCGCGCGGAGGCGGAGGACCTGGGCTCAGCGCGGTTACCTTGGCGCCCTCTATGGCAATG 606
 QY 606 AGGAGTCTGCGTGGGAGGATGCTCTGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 665
 DB 607 AGGAGTCTGCGTGGGAGGATGCTCTGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 666
 QY 666 CCAATGACCCCGGATAGTGTGCGCAACGTTGGTAAAGTCACTGATACCTTAACGTCGCG 725
 DB 667 CCACAGACCCCGGCGTGTAGTGTGCGCAATTTGGTAAAGTCACTGATACCTTACGTCGCG 726
 QY 726 GCTTTCGCGACCTCATGGGTTACATCCCTGTGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCG 785
 DB 727 GCTTTCGCGACCTCATGGGTTACATCCGCTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 786
 QY 786 GAGCTCTGCGCATGCGGTGAGGATCTCGGAGAGCGGGTTAATTTTGAACAGGGAAT 845
 DB 787 GGGCGCTGCGCATGCGGTGCGGCGGTTCTGGAAGACCGGCGTGAATATGCAACAGGGAAC 846
 QY 846 TACCGGTTGCTCTCTTTTCTATCTTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 905
 DB 847 TTTCTGTTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 906
 QY 906 TCTCCGCTGCGGAAGTGAAGAACATCAGTACCGGCTACATGGTGAATTAACGACTGCAACA 965
 DB 907 CTTTTCGCTTACCAAGTGGCAATTTCTCTCGGCGGCTTTTACCATGTACCAATGATTGCCCTA 966
 QY 966 ATGACAGCAATTACCTGGAGCTCAGGCTGCTCTCTCCAGCTCCCGGCGGCGGCGGCGGCGG 1025
 DB 967 ACTCGAGTATTTGTATCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1026

1026 GCGAAGAAAGTGGGAATGCATCTCAGTGTGGATACCGGTCTCACGAATGTGGCGGTGC 1085
 1027 GCGTTCCGAGAGGTAAAGCCCTCAGGTGTGGGTGGCGGTGACCCCAAGGTGGCACCA 1086
 1086 AGCGGCGCGCGCCCTCACGAGGGCTTGGGAGCGACATCGACATGTTGTGATGTCGG 1145
 1087 GGGAGGCAAACTCCCCACAACGAGCTTCGACGTATATCGATCTGCTTGTGGGAGCG 1146
 1146 CCAGCTCTGCTGCGCCCTTACGTGGGGACCTCTGCGGTGGGGTATGCTCGAGCCC 1205
 1147 CCACCTCTGCTGCGCCCTTACGTGGGGACCTGTCGGGTGCTCTTTCTTGTGGTC 1206
 1206 AATGTTTCAATGCTCGCCCGCAGCACCACTGGTTTGTCCAAAGCTGCAATTTGCTCATCT 1265
 1207 AACTGTTTCACTCTCTCCAGGGCACTGAGAGCGCAAGTCTGCAATTTGCTCATCT 1266
 1266 ACCCTGATCAATCACTGGAACACCGCATGATGGACATGATGAACTGTGCGCCA 1325
 1267 ATCCCGCCATATAACGGGTATCGCATGATGGGATATGATGAACTGTGCGCCA 1326
 1326 CGGCTACCATGATCTGGCGTACCGCATGCGTGTCCCGAGGTCAATATAGACATATTA 1385
 1327 CGGAGCGTTGGTGTAGCTCAGCTGCTCCGATCCCAAGGCATCATGGACATGATCG 1386
 1386 GCGGGCTCATTTGGGGCTCATGTTGGCTTGGCTTACTTCTATGACGGAGCGTGGG 1445
 1387 CTGGTCTCACTGGGAGTCTGCGGCGCATAGGTATTTCTCATGGTGGGAACTGGG 1446
 1446 CGAAGTCGTTGTATCATCTTCTGTTGGCCCGCGGGTGGACGCGCACCATCATCTGTTG 1505
 1447 CGAAGTCTCTGGTAGTCTGCTGCTATTTTCCGGGCTCGACGCGGAAACCCACGTCA 1506
 1506 GGGGTTCTGCGCGCAGACACCGGCGCTTACACAGCTTATTTGACATGGGCGCCAGCG 1565
 1507 GGGGAAATGCGCGCGCACCGCTGGCTGGCTTGTGCTCTCTTACACAGGCGCGCAAGC 1566
 1566 AGAAATCCAGCTCGTTAAACCAATGGCAGCTGGACATCAACCGCACCGCCCTCAACT 1625
 1567 AGAATCCCACTGATCAACCAACCGAGTTGGACATCAATAGCACGCGCTTGAAT 1626
 1626 GCAATGATCTTTGACACCGGCTTTATCGCTCTCTGTTCTACCCACAGCTTCAACT 1685
 1627 GCAATGAAGCTTAAACCGGCTGTTAGCAGGCTCTTCTATCAACAAATTTCAACT 1686
 1686 GGTGAGATGTCGGAACGATGTCGCTGCGCGAGTATCGAGGCTTCCGGGTGGAT 1745
 1687 CTTTCAAGCTGCTGAGAGTGGCCAGCTGCGGAGCGCTTACCGATTTTGGCCAGGGCT 1746
 1746 GGGGCGCTTTGCAATATGAGATAATGTCAACCAATCCAGAGGATATGAGACCTTATGCT 1805
 1747 GGGGCTCTATCAGTTATGCCAA-----CGAAGCGGCTTCGAGACGCGCTTCTGCT 1800
 1806 GGCACTACCAACCAAGCAGTGTGGGTGCTCTCGGAAAGATGTTGTGGCGCCAGTGT 1865
 1801 GGCACTACCTTCAAGACCTTGTGGCATTTGTCGCGCAAGAGCGTGTGTGGCGCGTAT 1860
 1866 ACTGTTTCAACCCCGCAGCTGTTAGTGGGACGACCGACAGCTTGGAGCGCCACTT 1925
 1861 ATTGCTTCACTCCCGCGCGGTGGTGGTGGGAACGACCGACAGTCTGGGCGCGCTACCT 1920
 1926 ACAGTGGGGAGAAATGAGACAGATGCTTCTTCTTATTTGAACAGCTTCGACACCGCTGG 1985
 1921 ACAGTGGGGTGCNAATGATAGGATGTTCTGCTTAAACACAGGCGCACCGCTGG 1980
 1986 GGTCAATGGTTCCGCTGACGCTGATGAATCTTTCTGGCTACACCAAGACTTGGGCGCAC 2045
 1981 GCAATTTGTTTGGTTGTTACCTGGATGAATCACTGGATTTCAACAAAGTGTGGAGCGC 2040
 2046 CACCTGCGGCTATGAGAGCTGATTTCAACCGCAGCAGGACCTGTTGTGCGCCACCGACT 2105
 2041 CCCCTTGTGTATCGGAGGGGTGGGCAACACA-----CCTTGTCTGTGCGCCACTGAT 2094

2106 GTTTTAGAAGCATCTGATACCACTTAACTCAAAATGCGGCTCTGGGCGCTGCTCACGC 2165
 2095 GCTTCGCGAACAATCGGAGGCAATATCTCTCGGTGCGGCTCCGCTCCCTGGATTACAC 2154
 2166 CAAAGTGTGATGATGATACCTTACAGGCTCTGGCATTTACCCCTGACAGTTAACTATA 2225
 2155 CCAGGTGATGATGATGATACCTTACAGGCTTGGCACTATCTTGTACCATCAATTACA 2214
 2226 CCATCTTCAAAATAGGATGATGAGGAGGGTGTAGACACAGGCTCACGGCTGATGCA 2285
 2215 CCATATTTCAAGTCAAGATGATGAGGAGGGTGTAGACACAGGCTGGAAGCGGCTGCA 2274
 2286 ATTTCACTCTGGGATCTGTCACCTTGGAGGACAGACAGAACTCAACTGTCTCTCTT 2345
 2275 ACTGAGCGGGGCGAAGCTGTGATCTGGAAGACAGGAGCTCCGAGCTCAGCCCT 2334
 2346 TGTGCACTTCCACCGAATGGGCAATTTTACCTTGTCTTATCTGGAAGCTTGGCGGCT 2405
 2335 TGTGCTGTCCACACAGTGGCAGGCTCTTCCGCTGTTCTTTCACGACCTGCCAGCT 2394
 2406 TGTGAGCTGCTCTTCCACCTCCACCAAAACATCGTGACGCTTACAAATTCATGTAGGC 2465
 2395 TGTCCACCGGCTCTATCCACCTCCACCAAGCAATTTGGACGTCGAGTACTTGTACGGG 2454
 2466 TATCACTCTCCCTCAAAATATACATCTCCGATGGAGTGGGTAATCTTATTTCTCTG 2525
 2455 TAGGCTCAAGCATCGGCTCTGGGCAATTAAGTGGGAGTACGTCGTTCTCTCTCTCT 2514
 2526 TCTTAGCGACGCGAGGTTTGGCTTGTCTTATGAGATGCTCATTTGTTGGGCGAGCGG 2585
 2515 TGTGTCAGACGCGGCTGCTCTCTGCTGTGATGATGTTACTCATATCCAAAGCG 2574
 2586 RAGCAGCTAGAGAGCTGCTTCTGCACTGCGAGCGAGCGAGCTAGCTGCAATTTGGCT 2645
 2575 AGCGGCTTTGGAGAACCTCGTAATCTCAATGACAGCTTCTTGGCGGAGCGACGGTC 2634
 2646 TCTATATTTGCTATCTTTTCTGCTGCTTGTGATCAATCAAGGCTCGGCTAGTCTCTCT 2705
 2635 TGTGCTCTCTCTGCTGCTTCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2694
 2706 TAGCTACTTATTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2765
 2695 GAGCGCTCTACGCGCTCTACGCGATGTCGCTCTCTCTGCTGCTGCTGCTGCTGCTG 2754
 2766 RAGCAGCTTATCTTATGACGCTCTGTCATGCGGAGCGAGCTAGCTGCTGCTGCTGCT 2825
 2755 AGCGGCTATACGCTGAGACGAGGCTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2814
 2826 TGATCACTCTCTTATCTCTCACCCCGGCTATAAGACCTTCTCAGCGGCTTTTGTGCT 2885
 2815 GGTAAATGGCGCTGACTCTGTCGCAATATTTACAGGCTATATCAGCTGTCATGCTG 2874
 2886 GGTGCTGCTATCTTCTGACCTTGGGGAGGCTATGCTCAGGAGTGGGACCACTATGCT 2945
 2875 GGTCTCAGTATTTCTGACAGAGTAAAGCGCACTGACGCTGCTGCTGCTGCTGCTGCT 2934
 2946 AGTGGCGGCTGCGCTGATGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3005
 2935 AGTGGCGGCGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2994
 3006 TGTGATCAATAACCAAGTGGCTTCTTGGCGGCTGCTTGGGCTGCTTACCTCTTAAAGG 3065
 2995 TATTTGATCATCAACAACTACTCTGCGCATCTTGGGCTTCTTGGATTTCTTCAAGCCA 3054
 3066 CTTTGAAGCGCTGCGCTTCTTCTGTCAGGCTTCAAGCTCTATGAGGATGTCGACCATGG 3125
 3055 GTTTGCTTAAAGTCCCTTCTGTCGCGCTTCAAGGCTTCTTGGGATGTCGCGCTAG 3114
 3126 CAGGCTCTGCGGCGGCGAGTACGCTCAGATGCGCTTACTAGCCCTTGGCAGTGA 3185
 3115 CCGGAGATAGCCGAGGCTTATGCTGCAATGGCCATCATCAAGTTTGGGCGGCTT 3174
 3186 CTGGCACTTATCATCTATGACCACTCACCCCTTATGCTGAGTGGGCTGCTAGTGGCTG 3245

QY 5406 GCTTGCACATTAACACGAGCGCGTCTGTCGCGCGGACAAAGAGGCTTCTTATGAGGCTT 5465
 DB 5395 GGATCGTCTTGTCCGGGAAGCGGCAATTATACCTGACAGGAGGTTCTCTACCAAGAGT 5454
 QY 5466 TTGATGAGATGGAGGATGTGCTCTAGGCGGCTCTATTGAAGAGGGGACGCGATAG 5525
 DB 5455 TCGATGAGATGGAGAGTGTCTCTCAGCACCTTACCGTACATCGAGCAAGGGATGATCTCG 5514
 QY 5526 CCGAGATGCTGAAGTCCAAGATCAAGGCTTATTGAGCAAGCTTCCAAACAGCTCAAG 5585
 DB 5515 CTGAGCAGTTCAAGCAGAGAGGCGCTCGGCTCTCTGAGACCGCTCCCGCATGCAAGG 5574
 QY 5586 ACATACAACCCACTGTGCGAGGCTTATGGGCCCAAGGTAGAACAAATTTGCGGCCAAACACA 5645
 DB 5575 TTATCACCCCTGCTCCAGACCACTGGCAGAACTCGAGGTCCTTTTGGGCGAAGCACA 5634
 QY 5646 TGTGGAACCTTATAGGCGATCAATATCCTCGAGGACTATCAACACTGCGCAAGGAAC 5705
 DB 5635 TGTGGAATTTTATCAGTGGGATACAAATCTTGGCGGCGCTGTCAAGGCTGCTGTGTAAC 5694
 QY 5706 CTCAGTGTGCTCCATGATGCGCTCAGTCCGCGCTCACCAGTCCGCTGTCAACAAGCA 5765
 DB 5695 CCGCCATTTGCTTATGATGGCTTTTACAGTCCGCTCACCAGCCCACTAACCACTGGCC 5754
 QY 5766 CCACTATCTTCTCAACATTTTGGGGGCTGGCTAGCATCCCAAAATTGCAACACCCGCG 5825
 DB 5755 AAACCTCTCTTCAACATATTGGGGGGTGGTGTGCTGCCAGCTCGCGCCCCGGTG 5814
 QY 5826 GGGCCACTGCTGCTTGTGTCAGTGGCTAGTGGAGCTGCGTAGGAGTATAGGCTTAG 5885
 DB 5815 CCGCTACTGCTGCTTGTGGGTGCTGGCTTAGCTGGCGCGCCATCGGCGAGCTTTGAGCTG 5874
 QY 5886 GTAAGTGTAGTGGACATCTCGCAGGATGATGTCGGGCAATTTTGGGGCTCTCGTGG 5945
 DB 5875 GGAAGTCTCTGTGGACATTTTGCAGGGTATGGCGGGGCTGGGGAGCTCTTTGTAG 5934
 QY 5946 CATTTAAGATCATGTCTGGCGAAGCCCTCCATGGAGGATGTGTCACACTTGTGCTGCTG 6005
 DB 5935 CATTTAAGATCATGAGCGGTGAGTCCCCCTCCACGAGGACCTGGTCAATCTGCTGCCG 5994
 QY 6006 GAATTCGTCTCGGGTGCCTTGGTAGTGGAGTCACTGCGCGGCGCATTTGCGCGGAC 6065
 DB 5995 CCATCTCTCGCTCGAGGCTTGTAGTGGTGTGCTGCGCAGCAATACTGCGCGCGC 6054
 QY 6066 AGTGGGACCGGGGAAGCGCGCTCCATGGATGATGACTCATTTGCTTGTGCTTCCA 6125
 DB 6055 AGTTTGGCCCGGCGAGGGGCGAGTGCATGGATGAACCGGCTTAATAGCCCTTGCCTCCC 6114
 QY 6126 GAGGAATCACGTGCGCCCCACCACTACGTGACGAGTGGATGCGTGCAGCGTGTGA 6185
 DB 6115 GGGGGAACCATGTTTCCCCACGCACTACGTGCGGAGAGCGATGACGCGCGCGGTCA 6174
 QY 6186 CCACACTACTTGGCTCCCTTACATAAACAGCTGTCTAGAGACTCCAACTGGAATTA 6245
 DB 6175 CTGCCATACTCAGCAGCCTCACTGTAAACCCAGCTCTGAGGCGACTGCATCAGTGGATAA 6234
 QY 6246 CTGAGACTGCCCATCCCATGCGGCGCTGCTGGCTCCGCGATGTGCGGACTGGTTT 6305
 DB 6235 GCTCGGAGTGTACCACTCCATGCTCCGGTTCTTGGCTAAGGACATCTGGGACTGGATAT 6294
 QY 6306 GCACCATCTTAAACAGACTTTAAATAATGGCTGACCTCCAAATTAATTTCCAAAGATGCCG 6365
 DB 6295 GCGAGGTGTGAGCGACTTTAAGACTGTGCTGAAGCGCAAGCTCATGCCAAGCTGCTG 6354
 QY 6366 GCTCCCTCTTGTCTCTGTCAAAAGGGGTACAAGGCGGTGTGGCGGCGACTTGGCATCA 6425
 DB 6355 GGAATTCCTTGTCTCTGCGAGCGGGGTATAGGGGGGTCTGGCGAGGAGACGGCATTA 6414
 QY 6426 TGACACACGGTGTCTTGGCGGCGCAATATCTTGGCAATGTCCGCTTGGGCTCCATGA 6485
 DB 6415 TGCACACTCGCTGCCACTGTGGAGCTGAGATCTCGACATGTCAAAAAACGGGACGATGA 6474

QY 6486 GAATCAACGGGGCTTAAGACCTGCATGAATATCTGCAGGGGACCTTTCTTATCAATTTGT 6545
 DB 6475 GGATCTCGTCTCTAGGACCTGACGAAACATGTGGAGTGGAGCTTCCCATTAACGCT 6534
 QY 6546 ACACGAGGGGCGAGTGGCGGAAACCGCGCGCAAACTTTAAGCTCGCATCTCGAGG 6605
 DB 6535 ACACACGAGGGCGGCTGTAACCTCCCTTCTGCGCGCAACTATTAAGTTGCGGCTGTGGAGG 6594
 QY 6606 TGGCGGCTCTCAGATACGCGAGGTGACGACGACGCGGTATACCACTACATAAACAGGAC 6665
 DB 6595 TGTCTGCAGAGGAAATACGTGGAGATAAGCGGGGTGGGAGCTTCCACTACGTATCGGTA 6654
 QY 6666 TCACTACATGATTAACCTGAAAGTCCCTGCAACTACCTCTCCGAGTTCTTTCTGCG 6725
 DB 6655 TCACTACTGACAACTCTTAAATGCGCGTCCAGATCCATCGCCCGAAATTTTTCACAGAAT 6714
 QY 6726 TGGACGGAGTGCAGATCCATAGTTTGGCCCAACCGGAGCGGTTTTCGCGGATGAG 6785
 DB 6715 TGGACGGGGTGGCGCTTACACAGGTTTGGCGCCCTTTCGAAAGCCCTTGTCTGCGGAGGAG 6774
 QY 6786 TCTCTGTTCTGCGTGGGCTTAATTCATTTGCTGCGGTCCAGCTTCTTCTGCGACCCCTG 6845
 DB 6775 TATCAATCAGAGTAGACTTCCAGATACCGGGTGGGTGCGAATTAACCTTGGAGCCCG 6834
 QY 6846 AACCCGACACAGACGTATTGATGTCCATGCTAACAGATCCATCTCATATCACGCGGAGA 6905
 DB 6835 AACCCGAGTAGCCGCTTGTGACGTCCATGCTCACTGATCCCTCCCATATAACAGCAGAG 6894
 QY 6906 CTGACGCGGGGTTTAGCGGGGGTCAACCCCATCCGAGGCAAGCTCTCTGCGGAGCC 6965
 DB 6895 CGCGCGGGAGAGGTTTGGCGAGAGGTTCAACCCCTTCTATGCGCAGCTCTCTGCGGTAGC 6954
 QY 6966 AGCTATCGGCACCATCTGCTGCGAGCACCTGCACCAACCCAGCGGCAAGCTTATGATGTG 7025
 DB 6955 AGCTGTGCTCATCTCTCAAGGCAATTGACCGCCCAACCATGACTCCCTCTGACGCGC 7014
 QY 7026 ACATGTTGATGCTAACTCT-----GTTTCATGGGGGCGATGTGACTCGGATAG 7073
 DB 7015 AGCTCATAGAGCTTAACCTCTGTCGAGGCGAGAGATGGGCGGCAACATCACAGGGTTG 7074
 QY 7074 AGTCTGGGTCCAAAGTGGTCTGTCGACTCTCTGACCCCATGTCGAAAGAGGAGCG 7133
 DB 7075 AGTCAGAGAACAAAGTGGTGTCTGAGCTCTCTGATCCGCTTGTGCGAGAGGAGATG 7134
 QY 7134 ACCTTTGAGCTTCTGATACCATCAGAAATACATGCTCTCCCAAGAAAGGTTCCACACAGCTT 7193
 DB 7135 AGCGGAGGTCTCCGTACTCTGCAAAATTTCTGCGAAGTCTCGAGATTTGCGCGGCGCC 7194
 QY 7194 TACCGGCTTGGGACGCGCTGATTACAAACCCAGCTTGTGGAATCTGTGAAAGGCGCAG 7253
 DB 7195 TGCCGCTCTGCGGCGCGGCGGACTACAAACCCCGCTAGTAGAGAGCTGGAAGGAGCTG 7254
 QY 7254 ATTACGACCGGCGACTGTTGGGGGTGTCCTCTCTCTCTTAGGAAACCCCGACGC 7313
 DB 7255 ACTAGCAACCACTGTGTGTCTTACGCTGCGGCTACCACTCTCCAGGCTCCCTCTCTGTG 7314
 QY 7314 CTCCTCCAGGAGGCGCGGACGAGTGGGCTTAAGTGGAGCTCCATAGGAGATGCCCTTC 7373
 DB 7315 CTCCTCCGGAAGAGCTACGTTGGTCTTCCAGGAATCAACCTTACTACTGCTTGG 7374
 QY 7374 AACAGTGGCCATTAAGTCTTTTGGCCAGCCCCCCCCCAAGCGGCGATTCAGGCTTTTCCA 7433
 DB 7375 CCGAGCTTGGCCACCAAAAGTTTGGCAGCTCTCAACTTTCGCGCATTTACGGGCGACAATA 7434
 QY 7434 CGGGGCGGCGGCTGCGGATTCGGGAGTGCAGACGCTCTCTGATGAGTTGGGCTTCTCGG 7493
 DB 7435 GCGAACACATCTCTG-----AGCCCGCCCTTCTTGGCTGCCCCCGCTCTCG 7482
 QY 7494 AGACAGGTTTCCATCTTTCATGCCCCCTCTGAGGCGGAGCTTGGAGATCCAGACCTGG 7553
 DB 7483 AGTTGAGTCTTATTTCTTCATGCCCCCTCTGAGGAGGAGCTTGGGATTCGGA----- 7537
 QY 7554 AGCTGAGCAGGTAGAGGCCCCCAACCCCCCCCCAGGGGGGTGGCAGCTCCCGGCTCG 7613

Db	7538	-----TCTCAGCGAG 7548	Db	8569	TCTGTAAAGTGGGGGTCAGAGGACGCGCGAGCCTGAGAGCTTTCACGGAGGCTA 8628
Qy	7614	ACTCGGGTCTCGGTCTACTTGTCTCGAGAGAGCACTCCGTCGTCGTGTCATGT 7673	Qy	8694	TCACAGGTATTTCTGCCCCCTCTGTCGACCCCCCAGACCGGAGTATGATCTTGAGCTGA 8753
Db	7549	GGTCATGTCGACGGTCACTAGTGTGGGGCCGACACGGAAGATGTCTGTCTCAATGT 7608	Db	8629	TCACAGGTACTTCGCGCCCCCGGGGACCCCCCAACACAGATACGACTTGAGACTTA 8688
Qy	7674	CATACTCTGGACGGGGCTTAATACTCTTGTAGTCCGGAAGAGGAGTATACCGA 7733	Qy	8754	TAAACATCTTCTCTCAATGTGTCTGTGGCGCTGGGGCCCAACAAGCCCGCCAGATACT 8813
Db	7609	CTTATTCTTGGACGGGCACTCGTCAACCCGTCGCTGGGAAGAACAAAACTGCCA 7668	Db	8689	TAAACATCATGCTCTCTCAAGTGTCTAGTCCGCCCAACGCGGCTGGAAGAGGCTTACT 8748
Qy	7734	TTAACCCCTTGAGCACTCCCTGTGCGATATCAACAAGGTGTACTGTACCAACAA 7793	Qy	8814	ACCTGACAGAGACCTACCACTCCAAATCGCCGGGCTGCTGGGAAACAGTTAGACACT 8873
Db	7669	TCACGCACTGAGCACTCGTGTCTAGCCATCACAACTGTGTATTCACCACTTCAC 7728	Db	8749	ACCTTACCGTACACCTCAACACCCCTCGCAGAGCGCGTGGGAGACAGCAAGACACA 8808
Qy	7794	AGAGCGCTCACTAAGGGCTTAAAGGTAACTTTGATAGGATGCAAGTCTCGACTCT 7853	Qy	8874	CCCTGTCAATTCATGGCTGGGAACATCATCTAGTACGCCCGCACCATATGGCTCGCA 8933
Db	7729	GCAGTGTCTGCCAAGGACAGAGAAAGTCACTTTGACAGACTGCAAGTTCTGGACAGCC 7788	Db	8809	CTCCAGTCAATTTCTGGCTAGGCAACAATAATCATGTTTGGCCCCACACTGTGGCGAGGA 8868
Qy	7854	ACTACGACTCAGTCTTAAAGGACATTAAGCTAGCGGCTCCAAAGGTCAACGCAAGCTCC 7913	Qy	8934	TGGTCTCTGATGACACACTTTCTTCTCATGCTCATGCTCAAGACACGCTGGACCAAGACC 8993
Db	7789	ATTACACAGGACGTCTCAAGAGGTCAAGCAGCGGCTCAAAAGTGAAGGCTTAATGTC 7848	Db	8869	TGATCTGATGACCCATTTCTTTAGGCTCTCATAGCCAGGATCAGCTTTGAACAGGCTC 8928
Qy	7914	TCACCATGGAGGCTTGCAGTTAAACCCCACTTCTGCAAGATCTTAAATATGGT 7973	Qy	8994	TCAACTTTGAGATGTACGAGCGGTGTACTCCGTGAGTCCCTTGGACCTCCAGCTATAA 9053
Db	7849	TATCCGTAGAGGAGCTTGCAGCTGAGCGCCCCCACTTCAGCCAAATCCAAAGTTGGCT 7908	Db	8929	TTAACTGTGAGATCTACGAGCGCTGTACTTCCATAGAACCACTGGATCTACCTCCAATCA 8988
Qy	7974	TTGGGGCTAAGGAGTCCGAGCTTGTTCGGGAGGCGGTAAACACATCAAGTCCGTGT 8033	Qy	9054	TTGMAAGGTTACATGGCTTGCCTTTTCTCTGCACATACACTCCCCACCAACTGA 9113
Db	7909	ATGGGGCAAGACGCTCGGTGTCATGCCAAGAGCGGTAGCCCATCAACTCCGTGT 7968	Db	8989	TTCAAGACTTCCATGGCTCAGCGCATTTTCACTCCACAGTTACTCTCCAGGTGAATCA 9048
Qy	8034	GGAAGACCTCTGGAGGCTCAGAACCAATCCCAACCAATATGGCCCAAAATG 8093	Qy	9114	CACGGTGTCTTTCAGCCCTCAGAAACTTTGGGGCGCCACCTCTCAGAGCGTGGAAAGCC 9173
Db	7969	GGAAGACCTTCTGGAGAGACGTGTAAACCAATAGACACTTACATCATGGCCCAAGACG 8028	Db	9049	ATAGGTGGCGGATGCTCAGAAACTTTGGGGTCCCGCCCTTGGAGCTGGAGACACC 9108
Qy	8094	AGGTGTCTCGGTGGACCCCAAGGGGGGCAAGAAAGCAGCTCGCTTATCGTTTACC 8153	Qy	9174	GGGCAGTGCAGTCAAGGCGTCCCTCATCTCCGTGGGGGAGAGCGCGTGTGGCTC 9233
Db	8029	AGGTGTCTCGGTGAGCTGAGAGGGGGGTGTAAGCCAGCTCGTCTCATCTGTCTCC 8088	Db	9109	GGGCGCGAGCGTCCGCGCTAGGCTTCTGTCCAGAGGAGGCGGCTGCATATGTGGCA 9168
Qy	8154	CTGACCTCGGCGTCAAGGCTGCGAGAGATGCGCTTTATGACATTAACAACAACTTC 8213	Qy	9234	GATATCTTCAATTTGGGGGTGAGACCAAGCTCAAACTCACTCCATTTGCCGAGCGC 9293
Db	8089	CCGACCTGGCGGTGCGGTGTGGAGAGATGCGCTGTACAGACGTGGTTAGCAAGCTCC 8148	Db	9169	AGTACCTCTTCAACTGGGCGAGTAAGAACAAAGCTCAAACTCACTCCAAATAGCGCGCGCTG 9228
Qy	8214	CTCAGCGGTGATGGGGCTTCTTATGATTCAGATTTCCCGCTCAGCGGTAGAGT 8273	Qy	9294	GGCTCTCTGATTTATCCAGCTGTTCACCGTGGCGCGCGCGGGGCGACATTTATCACA 9353
Db	8149	CCCTGGCGGTGATGGAGCTCTTACGATTCATTAATCACTCAGAGACAGCGGTTGAT 8208	Db	9229	GCGGCTGAGCTTGTCCGCTTGTTCACGCTGGCTTACAGCGGGGAGACATTTATCACA 9288
Qy	8274	TTCTCTTGAAGCATGGGCGGAAAGAGGACCTATGGGTTTTTGTATGATACCCGAT 8333	Qy	9354	GCGTGTGCGTGGCGGACCCCGCTTATGCTCTTTGGGCTACTCTACTTTTGTAGGGG 9413
Db	8209	TCCTGTGCAAGCTGGAAGTCCAAAGAACCCCGATGGGTTCTCGTATGATACCGCT 8268	Db	9289	GCGTGTCTCATGCGCGCGCGCTGGTCTCTGGTTTGGCTTACTCTCTGCTCGTGCAGGGG 9348
Qy	8334	GCTTTGATCAACGCTCAGTACAGAGACATCAGGACTGAGGATGCATATATCGGCGCT 8393	Qy	9414	TAGGCGTTTTCTTACTCCCGCTCGGTAGAGCGGACACATTAGCTTACACTCCATAGCTA 9473
Db	8269	GTTTGTACTCCAGTCACTGAGAGGACATCCGTACGGAGAGGCAATTTACCAATGTT 8328	Db	9349	TAGGCACTACCTCTCTCCCAACCGAGAGGTTGGGGTAAACACTCCGCGCTCTTAAGC 9408
Qy	8394	GCTCTTGGCGGAGGCGCCACACTGCCATACACTCGCTAACTGAGAGACTTTACGTGG 8453	Qy	9474	ACTGTCCCTT 9533
Db	8329	GTAACCTGGAACCCCAAGCCCGGTGGCCATCAAGTCCCTCACTGAGAGGCTTTATGTTG 8388	Db	9409	CATTTCTCTT 9468
Qy	8454	GAGGGCTATGTTCAACAGCAAGGGCCAAACCTGCGGGTACAGCGGTTGCGCGCCAGCG 8513	Qy	9534	TT 9593
Db	8389	GGGGCTCTTACCAATTAAGGGGGAAACTCGCGGTACCGCAGGTGCGCGCGAGCG 8448	Db	9469	TTCTTTTTTCT 9485
Qy	8514	GGGTGCTCAACACTAGCATGGGAAACCAATCACTAGCTACGTAAGGAAAGCTTACGGGCTT 8573	Qy	9594	CTTATTTACTTTCTTTCT 9653
Db	8449	GCSTACTGACAACTAGCTGTGGTAAACACCCCTCACTTGTCTACATCAAGGGCCCGGCGACCT 8508	Db	9486	-----TTTTTCTTTTAAATGGTGGCTTCCATCTTAGCCCCCTAGTACGGCTAGCTGAAA 9541
Qy	8574	GTAAGCTGAGGAGTAAATCGCGCCCAATGCTGTATGCGGAGTACTTCTGTGTGCA 8633	Qy	9654	GCTCCGTGAGCGCGCTAGCTGAGAGAGTCCGCTTAACTGCTCTCTCTGAGATCATGT 9711
Db	8509	GTCAGCGCGAGGCTCCAGGACTGACCACTGCTGCTGTGTGGCGAGCACTTATGTCGTTA 8568	Db	9542	GCTCCGTGAGCGCGCTAGCTGAGAGAGTCTGATACTGCGCTCTCTCTGAGATCATGT 9599
Qy	8634	TCTCAGAAAGCCAGGGGACCGAGGAGACGAGCGGAACCTGAGAGCCTTTCACGGAGGCTA 8693			

```

; Sequence 4, Application US/09014416
; Patent No. 6153421
; GENERAL INFORMATION:
; APPLICANT: Yanagi, Masayuki
; APPLICANT: Bukh, Jens
; APPLICANT: Emerson, Susanne U.
; APPLICANT: Purcell, Robert H.
; TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 20264276
; CURRENT APPLICATION NUMBER: US/09/014,416
; CURRENT FILING DATE: 1998-01-27
; EARLIER APPLICATION NUMBER: US 60/053,062
; EARLIER FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 9595
; TYPE: DNA
; ORGANISM: Hepatitis C virus
; US-09-014-416-4

Query Match          46.4%;   Score 4505,8;   DB 3;   Length 9595;
Best Local Similarity 67.3%;   Pred. No. 0;
Matches 6542;   Conservative 0;   Mismatches 3032;   Indels 147;   Gaps 7;

6  CCCCTAATAGGGGCGACACTCCGCGCATGAATCACTCCCTGTGAGGAACTACTGTCTTCA 65
7  CCCCTGATGGGGCGACACTCCACCATGAATCACTCCCTGTGAGGAACTACTGTCTTCA 66

66  CGCAGAAAGCGTTAGCCATGGCGTTAGTATAGTGTGTGTA CAGCTTCAGGCCCCCCCCC 125
67  CGCAGAAAGCGTTAGCCATGGCGTTAGTATAGTGTGTGTA CAGCTTCAGGCCCCCCCCC 126

126  TCCCGGAGAGCCATAGTGTCTCGCGAACCGGTGAGTACACCGGAATTCGCGGAAAGAC 185
127  TCCCGGAGAGCCATAGTGTCTCGCGAACCGGTGAGTACACCGGAATTCGCGGAGACAC 186

186  TGGGTCTCTTCTTGGAATAAACCACTATGCCCGGCCATTTGGGGCGTGCCCGCAAGA 245
187  CGGCTCTTCTTGGATCAACCGCTCAATGCTTGAGATTTGGGGGTGCCCGCGAGA 246

246  CTGCTAGCCGAGTAGCGTTGGGTTGCGAAAGCGCTTGTGGTACTGCTGATAGGGTGCTT 305
247  CTGCTAGCCGAGTAGTGTGGGTCGCGAAAGCGCTTGTGGTACTGCTGATAGGGTGCTT 306

306  GCGAGTCCCGCGGGAGGTCTCGTAGACCGTGACCATGAGCAAAATCCTAAACCTCAA 365
307  GCGAGTCCCGCGGGAGGTCTCGTAGACCGTGACCATGAGCAAAATCCTAAACCTCAA 366

366  GAAAAACCAAAAGAAAAACAACACCGTCGCCCAACAAGACGTAAAGTTTCGGGCGCGCGCC 425
367  GAAAAACCAAAAGCTAAACACCAACCGCGGCCCAAGAGCGTCAAGTTCCCGGGCGGTGTC 426

426  AGATCGTTGGCGGAGTATACCTGTTGCGCGCAGGGGCCCAAGTTGGGTGTGCGCGCGA 485
427  AGATCGTTGGTGGAGTTTACCTGTTGCGCGCGAGGGGCCCAAGTTGGGTGTGCGCGCGA 486

486  CAAGGAAGACTTTCGAGCGGTCCAGGCACGTGGAAGGGCCAGCCCATCCCTAAAGATC 545
487  CTAGGAAGGCTTTCGAGCGGTTCGCAACTCTGTTGGAAGGGGCAACCATATCCCAAGGCTC 546

546  GGCGCTCCACTGGCAAAATCTGCGGGAACACAGGATACCCCTGGGCCCTATACGGGAATG 605
547  GCCACCCGAGGGCAGGGCCCTGGGCTCAGCCCGGGTACCTTGGGCCCTCTATGGAATG 606

606  AGGGAATCTGGGCGCAGGATGGCTCTGTTCCCGCGAGGTTCCTCGCTCTTGGGGGCC 665
607  AGGCGCTGGGTGGCAGGATGGCTCTGTGTCAACCCCGCGCTCCCGGCTAGTTGGGGCC 666

666  CCAATGACCCCGCGCATAGGTTCGCGCAAGCTGGGTAAAGTCAATCGATACCCCTAAAGTGG 725
667  CCAAGGACCCCGCGGTAGGTTCGCTAACTTGGGTAAAGTCAATCGATACCCCTTAATGCG 726

```


Db 3955 TCGGGTCTCGGCTTTACAGACAACTCAACCCCGCGCTGTACCGCAGACATTTCCAAAG 4014
 QY 4026 TCGGGTACTTGGATCGCCCGGAGTGGCAGTGGAAAGAGCACCAGAAAGTTCTGTGCGCATATG 4085
 Db 4015 TGGCAGATCTGCAGGCTCTACTTGGCAGCGGCGAGAGCACCAGAGTGGCGGCTGCGTATG 4074
 QY 4086 CTGCTCAGGGGTATAAGTGTAGTGTCTTAATCCCTCAGTGGGTGCCACCTCGGGGTTTG 4145
 Db 4075 CAGCCCAAGGGTACAAGGTGCTCGTCTGAACCGGTCCGTTGCCGCCACCTTAGGGGTTTG 4134
 QY 4146 GGGGCTACTTGTCTAAGGCACATGGGCATCAATCCCAACATTAGGACTGGAGTCAGACTG 4205
 Db 4135 GGGCGTATATGTCCAAAGGCACACGGTATCGACCTTAACATCAGAACTGGGGTAAAGACCA 4194
 QY 4206 TGACGACCGGGCGGCCCATCAGCTACTCCACATATGCGCAATTCCTCGCGGATGGGGGT 4265
 Db 4195 TTACACAGGGCGGCTCAATTAAGTACTCCACTATGCGCAAGTTCCTTGGCGAGGTGGCT 4254
 QY 4266 GTGCGGGCGGCTTACGACATCATATGTGATGAATGCCATGCCGTGGACTCTPACCA 4325
 Db 4255 GTTCTGGGGCGGCTATGACATCATATATGTGATGAGTGGCACTCAACTGACTCGACTA 4314
 QY 4326 CCATCTTGGCATCGGAAAGTCTTGTATCAAGCAGACAGCTGGGGTCAAGTCAACTG 4385
 Db 4315 CCATCTTGGGCATCGGCACAGTCTTGACCAAGCGGAGCGGCTGGAGCGCGGCTCGTCG 4374
 QY 4386 TGTGCTACAGCTACGCGCCCTGGGTGAGTGAACACCCCGCCCAACATAGAGGAGG 4445
 Db 4375 TGTCTGCCACCGCTACACTCGGGATCGGTTACGTGCCACACCCCAATATCGAGGAAA 4434
 QY 4446 TGGCCCTTGGGCAAGGAGGCGAGATCCCTTCTATCGGAGGGCGATTTCCCTGTCTTACA 4505
 Db 4435 TAGGCCCTGTCAAACATAGGAGATCCCTTCTATGGCAAGCCATCCCATTTAGAGGCCA 4494
 QY 4506 TCAAGGAGGAGACATCTGATCTTCTGCCATTCAAAGAAAGTGTGACGAGCTCGCG 4565
 Db 4495 TCAAGGGGGAGGACATCTATTTCTGCGCATTTCCAAAGAAAGTGTGACGAGCTCGCG 4554
 QY 4566 CGGCGCTTCGGGTATGGGCTTGAATCTAGTGCATCTACAGAGGTTGGAGCTCTCG 4625
 Db 4555 CAAGCTGACAGGCTCGGACTGACGCTGACGCTGAGCATATTTACCGGGGCTTGTGTCTCG 4614
 QY 4626 TAATACCAACTCAGGAGAGCTAGTGTGTGCTGCCACCGACGCTCATGACAGGGTATA 4685
 Db 4615 TCATACCGCTATCGGAGAGCTGTTGTGTGCGCAACAGAGCTCTAATAGAGGGTTCA 4674
 QY 4686 CTGGGAGCTTGTGCTCGGTATCGACTGCAAGCTAGCGGTCTCAAGTGTGAGACTTCA 4745
 Db 4675 CCGGCGATTTGACTCAGTGTATCGACTGCAATACATGTGTCAACCCAGACAGTCTGACTTCA 4734
 QY 4746 GTTTAGACCCCATTCACCATAAACACACAGATTTGCTCCTCAAGACGCTGTCTACGTA 4805
 Db 4735 GCTTGTATCCACTTTCACCATTTAGACAGACGACCGTGGCCCAAGACGCGGTGTGCGCT 4794
 QY 4806 GCAGGCGCGGGTTCGACCGGTAGGGAGAGCTGGGCAATTTATAGTATGTTTCCACTG 4865
 Db 4795 CGCAAACGGGAGTGAAGTGGCAGGGTAGGAGTGGCATCTACAGTTTGTGACTCCAG 4854
 QY 4866 GTGAGCGAGCTCAGGAATGTTTGAAGTGTAGTGTCTGTGAGTGTGACGAGCGAGGG 4925
 Db 4855 GAGAAGCGGCTTCGGGATGTTTGGATTTCTCGGTCTGTGTGAGTGTGATGAGCGGGCT 4914
 QY 4926 CCGCATGTTAGTGTACACCATCGGAGACACCGTCAAGGCTCAGGCTCAGGGGTATTTCAACA 4985
 Db 4915 GTGCTTGGTATGAGTGTACCGCGCTGAGACCTTCGGTTAGTTTGGGGCTTACCTAAATA 4974
 QY 4986 CGCCCGGTTTGGCTGTGTGCAAGACCATCTTGAAGTTTGGGAGGCAATTTTTCACCGGG 5045
 Db 4975 CACCAGGGTTGCGGCTGCGCAGGACCATCTGAGTTTCTGGAGAGCGTCTTTCACAGGCC 5034
 QY 5046 TCACACACATAGATGCCACTTCTTTCCTTCCCAACAAAGCATCGGGGGGAAATTTTCGAT 5105
 Db 5035 TCACCCACATAGATGCCACTTCTTCTGTCCAGACTAAACAGGCGAGGAGCAACTTTTCCTT 5094

QY 5106 ACTTAACAGCCTTACAGGCTACAGTGTGCTAGGGCCAAAGCCCCCCCCCGTCTCTGG 5165
 Db 5095 ACTTGGTGGCATATCAAGCTACAGTGTGCGCCAGGCTCAAGCTCCACCTCCATCGTGG 5154
 QY 5166 AGTCATGTGGAAGTGTGTGACTGACTCAAGCCACACTCGTGGGCCCCACACTCTCC 5225
 Db 5155 ACCAAATGTGGAAGTGTCTCATACGCTGAAACCTACACTGACGGGCCAACACCCCTGC 5214
 QY 5226 TGTACCGCTTGGGCTGTGTTACCAACAGGCTCACCTCACATCCCGTACGAAATACA 5285
 Db 5215 TGTATAGGCTAGAGCGCTCCAAATAGGTCATCTCTCACACCCCATAACTAAATACA 5274
 QY 5286 TCGCCACCTGCAATGAAGCCGACCTTGAAGTCAATCACCAGCAGTGGGTCTTGGCAGGG 5345
 Db 5275 TCATGGCATGATGTGCGCTGACCTGAGGTCTGCTACTAGCAGCTGGGTGCTGGTAGCG 5334
 QY 5346 GAGTCTTGGCGGCGTTCGCGCGTATGTCTGCGGACCGGGTGTGTTGCATCATCGGCC 5405
 Db 5335 GAGTCTTGCAGCTTGTGCGCGCATACTGCTGACGACAGGCAAGTGTGCTATTGTGGGCA 5394
 QY 5406 GCTTGCACATTAACAGCGAGCGGTGCTGCGCGGACAGAGGCTCTCTATGAGGCTT 5465
 Db 5395 GGATCATCTTGTCCGGGAAGCCAGCTGTGTTCCGACAGGGAAGTCTCTTACCAGAGT 5454
 QY 5466 TTGATGAGATGAGGAATGTGCTCTTAGGGCGGCTCTATTGAAGAGGGCGAGCGGATAG 5525
 Db 5455 TCGATGAGATGGAAGTGTGCTTCAACATTTCTTACATCGACGAGGAATGACAGCTCG 5514
 QY 5526 CCGAGATGTGAAGTCCAAAGTCCAAAGCTTATTTGAGCAGCAAGCTTCCAAAGCTCAAG 5585
 Db 5515 CCGAGCAATTCAGCAAAAGGCGCTCGGGTGTGTTGCAACGCGCCACCAAGCAAGCGGAG 5574
 QY 5586 ACATACAAACCCACTGTGAGGCTTCATGSCCCCAAGGTAGAACAAATTTCTGGGCCAACA 5645
 Db 5575 CTGCTGCTCCCTGGTGGAGTCCAAGTGGCGAGCCCTTGAGACCTTTCTGGGCGAAGACA 5634
 QY 5646 TGTGAACTTCAATTAGCGGCATCCAAATACCTCGCAGGACTATCAACACTGCCAGGGAACC 5705
 Db 5635 TGTGGAATTTTACATCAGCGGAATACAGTACTAGCAGGCTTATCCACTCTGCTCGGAAACC 5694
 QY 5706 CTGAGTGTCTTCAATGATGAGGCTTCAAGTCCGCGCTTACCAAGTCTGCTGTCAAGCA 5765
 Db 5695 CCGCGATAGCATCAATGATGAGCAATTTACAGCTTCTATCACTAGCCGCTCAGCACCCAAA 5754
 QY 5766 CCATCTCTTCTCAACATTTTGGGGGCTGCTAGCATCCCAATGCAACCCACCCGCG 5825
 Db 5755 ACACCTCTCTGTTAAACATCTTTGGGGGAGTGGTGGCTGCCCACTCTCTCCAGCG 5814
 QY 5826 GGGCCACTGGCTTCTGTGTGCTAGTGGGCTAGTGGGAGCTGCCGTAGGCACTATAGGCTTAG 5885
 Db 5815 CTGCGTACGTTTCTGTGGGCGCGGATCGCGGAGCGGCTGTTGGCAGCATAGGCTTGG 5874
 QY 5886 GTAAGTGTCTAGTGACATCTTGGCAGGGTATGTTGGCGGCAATTTGGGGGCTCTCGTGG 5945
 Db 5875 GGAAGTGTCTGTGACATCTTTGGCGGCTATGGGCGAGGGGTAGCGCGGCACTCGTGG 5934
 QY 5946 CATTCAGATCATGTCTGGGAGAGCCCTCCATGGAGATGCTGCTCAACTTGTCTGCTG 6005
 Db 5935 CTTTAAAGTCAATGAGCGGCGAGGTCCTTCCACGAGGACCTGGTCAACTTACTCTCTG 5994
 QY 6006 GAATTTCTGTCTCGGCTGCTTGGTAGTGGGAGTCACTGCGGCGCATTTCTGCGCGGAC 6065
 Db 5995 CCATCTCTCTCTGTCGCTGGTCTGCTGGGCTGCTGCGGCAATACTGCGTGGC 6054
 QY 6066 AGTGGGACCGGGGAGCGCGCTCCAAATGATGAATAGATCTCATTTGCTTTGCTTCCA 6125
 Db 6055 AGTGGGCGCGGAGAGGGGGCTGTGCAAGTGAATGAACCGGCTGATGAGCTTCTGCTTCCG 6114
 QY 6126 GAGGAATCAGCTCGCCCCCAGCTAGTGAACGAGTTCGGATGCTGCGAGCGGTGCA 6185
 Db 6115 GGGGTAACCACTCTCTCCCTCATGCTATGTGCTGAGAGCGAGCTGCGAGCAGGTGCTCA 6174

	7255	ACTACGTCCTCCCGTGGTA	CACGGATGCCCATTTGCACACCTACCAGGCTCCTCCAAATAC	7314
Ddb				
	7314	CTCCCCAAGAGGCGCGACAGTGGGCGCTTAAGTGAGGAATCCCATAGGAGATGGCCCTTC	7373	
Qy				
Ddb		7315	CACCTCACGGAGAAAAGAGACGGTGTCTCTGACAGAATCCAATGTGTCTCTCTGCTGG	7374
	7374	AACAGCTGGCCATTAAAGTCTTTTGGCCAGCCCCCCCAGCGGGGATTCAGGCGCTTTCCA	7433	
Qy				
Ddb		7375	CGGAGCTCGCCACTAAGACCTTCGG-----TAGCT	7404
	7434	CGGGGCGGGCGCTGCCGATTCGGCAATTCGGCAGTCAGACGCTCCTGATGAGTTGGCCCTTTCCG	7493	
Qy				
Ddb		7405	CCGATCGTCCGCCGTTGATAGCGCAGCGCGACCGCCCTTCTTGACCTGGCCCTCCGACG	7464
	7494	AGACAGGTTCCATCTCTTCCATGCCCCCTCGAGGGGGAGCTTGAGATCCAGACCTGG	7553	
Qy				
Ddb		7465	AC-----GGTGACAAAGGATCCG	7482
	7554	AGCCTGAGCAGGTAGAGCCCCAACCCCCCCCCAGGGGGGTGGACGCTCCCGGCTCGG	7613	
Qy				
Ddb		7483	ACGTTGAGTCGTACTCTCCATGCCCCCTTGAGAGGGAGCGGGGACCCCGATCTCA	7542
	7614	ACTCGGGTCTGGCTTACTTTGCTCCGAGAGG---ACGACTCCGTCTGTGCTGCTCCA	7670	
Qy				
Ddb		7543	GCGACGGGCTTGTGCTACCGTGAGTAGGAGGCTAGTGAGGATGTCTGTGCTGCTCAA	7602
	7671	TGTCATACTCCTGGAACGGGCTCTAATAAATCTCTTGTAGTCCCGAAGAGAGAGTTAC	7730	
Qy				
Ddb		7603	TGTCCTATACGTGAGCAGGCGCCTGTATACGCCATGCGCTGCGGAGNAAGTAAGCTGC	7662
	7731	CGATTAAACCCCTTGAGCAACTCCCTGTTGCGATATACAACAAGGTGACTGTACCAAA	7790	
Qy				
Ddb		7663	CCATCAACCCGTTGAGCAACTCTTTGTGCGTCAACCACAAATGCTACGCCACAACAT	7722
	7791	CAAAAGAGCGCTCAATAAGGCGTAAAAAGGTAACTTTTGATAGATGCAAGTGTCTGACT	7850	
Qy				
Ddb		7723	CCCGCAGCGCAAGCCTCCGGCAGAAAGGTCACTTTTGACAGATGTGCAAGTCTCGGATG	7782
	7851	CCTACTACGACTCAGTCTTAAGGACATTAAGCTAGCGGCGCTCCAAGGTACCGCAAGGC	7910	
Qy				
Ddb		7783	ATCATTTACCGGACGTACTCAAGGAGATGAAGCGGAGCGGTCCACAGTTAAGCTAAGC	7842
	7911	TCCTCACCATGGAGGAGGCTTGCCAGTTTAACCCCAACCCCATTTCTGCAAGATCTAAATATG	7970	
Qy				
Ddb		7843	TTCTATCTATAGAGGAGGCTGCAAGCTGACGCCCCACATTTGCGSCCAATCCAAATTTG	7902
	7971	GGTTTTGGGCTAAGAGGTCCGAGCTTGTTCGGGAGGGCGGTTAAACACATCAAGTCCG	8030	
Qy				
Ddb		7903	GCTATGGGGCAAAAGGACGTCGCGAAACCTTATCCAGCAGGGCGGTTAAACACATCCGCTCG	7962
	8031	TGTGGAAGGACCTCCTCGAGGACTCAGAAAAACCAATTTCCACAAACCATTTATGGCCAAA	8090	
Qy				
Ddb		7963	TGTGGAGGACTTGTCTGGAAGACCTGAACACCAATTTGACACCACTATGCGCANAAA	8022
	8091	ATGAGTGTCTTGCTGTGACCCCAACAAAGGGGGCAAGAAAGCAGCTCGCCCTTATCGTTT	8150	
Qy				
Ddb		8023	GTGAGTTTTCTGCTGCCAACAGAGAGGGAGGCGCGAAGCCAGCTCGCCCTTATCGTAT	8082
	8151	ACCTGACCTCGGCGTCAAGGTCTGCGAAGAGATGGGCCCTTTTATGACATTTACACAAAAC	8210	
Qy				
Ddb		8083	TCCAGACCTGGGAGTTCTGTATGCGAAGAGATGGCGCTTTTACGACGTGTCTCACCC	8142
	8211	TTCCTCAGGCGGTGATGGGGCTTCTTATGATTCAGTATTTCCCGCTCAGCGGGTAG	8270	
Qy				
Ddb		8143	TTCTCTCAGGCGGTGATGGGCTCTCTATACGGATTTCAATATCTCCCCCAAGCAGCGGTCG	8202
	8271	AGTTTCTCTGAAAGCATGGGCGGAAAAAGAGACCTTATGGTTTTTTCGTATGATACCC	8330	
Qy				
Ddb		8203	AGTTCTCTGGTGAATACCTGGAAATCAAGAAATGCCCTATGGGCTTCTCATATGACACC	8262
	8331	GATGCTTTGACTCAACCGTCTACTGAGAGAGACATCAGGACTTGAGGAGTCCATATATCGGG	8390	
Qy				

[illegible]

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```

	Query Match	45.6%	Score 4428.6;	DB 5;	Length 9401;
	Best Local Similarity	67.2%;	Pred. No. 0;		
	Matches 6351;	Conservative	20;	Mismatches 2990;	Indels 90; Gaps 5;
Qy	6	CCCCTAATAGGGGGCAGACTCCGCCATGAATCACTCCCTGTGAGGAAC	ACTCTGCTTCA	65	
Db	7	CCCCGTGATGGGGCAGACTCCACCATGAATCACTCCCTGTGAGGAAC	ACTCTGCTTCA	66	

QY	66	CGCAGAAAGCGTCTAGCCATAGCGGTTAGTATGAGTGTCTGTAAGGCTTCAGGCCCCCCCC	125
Db	67	CGCAGAAAGCGTCTAGCCATAGCGGTTAGTATGAGTGTCTGAGGCTCCAGGACCCCCC	126
QY	126	TCCCGGAGAGCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTCGCGGAGAC	185
Db	127	TCCCGGAGAGCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTCGCGGAGAC	186
QY	186	TGGGTCTCTTCTTTGGATAAACCACCTATATGCCCGGCCATTTGGGGCTGCCCGCAAGA	245
Db	187	CGGGTCTCTTCTTTGGATCAACCCGCTCAATGCTTGAGATTTTGGGCGTCCCGCGCAAGA	246
QY	246	CTGCTAGCGAGTAGCGGTTGGGTTCGAAAGGCGCTTGTGGTACTGCTCTGATAGGGTGCTT	305
Db	247	CTGCTAGCGAGTAGGTTGGGTTCGAAAGGCGCTTGTGGTACTGCTCTGATAGGGTGCTT	306
QY	306	GGAGTGCCTCCGGAGGTTCTGTPAGACGCTGACCAATGAGCAAAATCTCTAAACCTCAAA	365
Db	307	GGAGTGCCTCCGGAGGTTCTGTPAGACGCTGACCAATGAGCAAAATCTCTAAACCTCAAA	366
QY	366	GAAAAACCAAGAAACACCAACCGTTCGCCACAGACGTTAAGTTTCGGGGGGGGCC	425
Db	367	RAAAAAACCAAGAAACACCAACCGTTCGCCACAGACGTTAAGTTTCGGGGGGGGCC	426
QY	426	AGATCGTTGCGGAGTATCTCTGTTCCCGCGCAGGGGCCCAAGTTTGGGTGGCGCGCA	485
Db	427	AGATCGTTGTTGAGTATCTCTGTTCCCGCGCAGGGGCCCTAGATTGGGTGGCGCGCA	486
QY	486	CAAGGAAGACTTCCGGAGCGTCCACGCCAGTGGAAAGCGCCAGCCCATCCCTAAAGATC	545
Db	487	CGAGAAAGACTTCCGAGCGGTGCAACCTCGAGGTAGACGTGAGCCTATCCCCAAGGCTC	546
QY	546	GGCGCTCCACTGCAAAATCCTGGGGAAAAACCGAGTACCCCTGGGCCCTATACGGGAATG	605
Db	547	GTCGGCCCGAGGCGCAGGACCTTGAGGCTCAGCCCGGGTACCTTTGGGCCCTCTATGGCAATG	606
QY	606	AGGGACTTCGCTCGGCAGAGTGGCTCTGTCGCCCGCAGGTTCCCGTCCCTCTTGGGGCC	665
Db	607	AGGGCTCGGGTGGGCGGAGTGGCTCTGTCGCCCGTGGCTCTCGGCTAGCTGGGGCC	666
QY	666	CCAATGACCCCCCGGCATAGTTCGCGCAACGTGGGTAAAGTTCATCGATACCTTAACGTGG	725
Db	667	CCACAGACCCCCGCGTAGTTCGCGCAATTTGGGTAAAGTTCATCGATACCTTACGTGG	726
QY	726	GCTTTGCGACCTCATGSGGTAATCCCTGTCTGTGGGGCCCGGCTCGGGGGGTGCGCA	785
Db	727	GCTTCGCGACCTCATGSGGTAATCCCTGTCTGTGGGGCCCGGCTTCGTGGAGGCGCTGCGCA	786
QY	786	GAGCTCTCGGCATGGGTGAGAGTCTCGGAGACGGGTAAATTTGCAACAGGGAAT	845
Db	787	GGGCCCTGGGCGATGGCGTCCGGGTTCTGGAACCGCGTGAATATGCAACAGGGAATC	846
QY	846	TACCGGTTGCTCCTTTTCTATCTTTGCTGGCCCTGCTGCTGCATCACACCCCGG	905
Db	847	TTCTGGTTGCTTTCTTCTAYCTTCTCTGGCCCTGCTCTCTTGCTTGACTGTGCGCG	906
QY	906	TCTCCGCTGCCAAGTGAAGAACATCAGTACCGGCTACATGGTGACTAAAGTCTGACCA	965
Db	907	CTTCGGCCTACCAAGTGGCAATCTCCACGGGGCTTTACCAACGTCAACATGATTGGCCTA	966
QY	966	ATGACAGCATTAATCTGGCAGCTCAGGCTGCTGCTCTCCACGTCCCGGGTGGTCCCGT	1025
Db	967	ACTCGAGTATGTGTACGAGGGCGCGGATGCCATCCTGCACACTCCGGGGTGGTCCCTT	1026
QY	1026	GCGAGAAAGTGGGGAATGATCTCAGTGTGGATACGGGTCTCACCAAAATGTGGCGGTGC	1085
Db	1027	CGGTTGTTAGGGCAACCGCCTCGAGGTGTTGGGTGGGATGACCCCTACGTTGGCCACCA	1086
QY	1086	AGCGGGCCGGCCCTCACGAGGGCTTCGGGACGCAATCGACATGGTTGTGATGTCG	1145
Db	1087	GGGATGGCAACTCCCCCGACACGAGCTTCGACGTCAATCGATCTGCTGTGCGGAGCG	1146

3295	CGTGGGGGAGATACCGCCCGCTGCGGTGACATCATCAACGGCTTGCTGTTCTTCGCGCC	3354	3354
3366	GACTTGGTTCGGGAGGTCTCTCTTGGCCAGCTGATGGCTATACCTCCAAAGGGGTGGAGTC	3425	3425
3355	GCAGGGGCGGGAGATACTGCTCGGGCCAGCCGATGGAAATGGTCTCTCAAGGGGTGAGGT	3414	3414
3426	TTCTCGCCCCCATCATCTGCTTTACGCCCAGCAGACACGTGGGCTTTTGGGACCATAGTGG	3485	3485
3415	TGCTGGCGCCCATCACGGCGTACGCCCAGCAGACAAGGGGCTCTCATGGGTGCATAATCA	3474	3474
3486	TGACATGACGGGGCGGACAAAGACAGAAACAGGCTTGGGAAATTCAGGTCTGTCCACAG	3545	3545
3475	CCAGCTTACTTGGCCGGGACAAAACCAAGTGGAGGTGAGGTTCGAAATTTGTGTCACTG	3534	3534
3546	TACTCAGTCTCTCTCTCGGAACATCCATCTCGGGGTTTTTGTGGACTGTCTACCATGGAG	3605	3605
3535	CTGCCCAACCTTCTCTGCAACGTGCATCAATGGGTGCTGTGGACTGTCTACCAAGGG	3594	3594
3606	CTGGGAACAAGACTCTGGCCGGCTCA CGGGGTTCGGTCA CGGAGATGTACTCCAGTGTG	3665	3665
3595	CCGAAACAGGACCATTCGCTGTCACCAAGGCTCTGTCTATCCAGATGTATACCAATGTAG	3654	3654
3666	AGGGGGACTTAGTAGGGTGGCCAGCCCTCGGGACTTAATCTTTTGAGCGCTGCACGT	3725	3725
3655	ACCAAGACCTTTGTGGGCTGGCCCGCTYCGCAAGGTATCCGGTCTATGACACCTGCACCT	3714	3714
3726	GTGGAGCGGTCCGACTGTACTCTGCTCAGCGGGAACGCTGATGTCTATCCGGTTCGAAGAC	3785	3785
3715	GGGCTCTCTCGAACCTTTACTCTGCTCAGGAGCAGCCGATGTCTATCCCGTGC CGCG	3774	3774
3786	CGGGGGAACAACGGGAGCGTACTCTCTCCCGAGACTCTTTTCCACTTTGAAGGGGTCTCT	3845	3845
3775	GGGGTGATAGCAGGGGAGCGCTGCTGTGCGCCCGGCCCATTTTCTTACTTTGAAGGCTCTCT	3834	3834
3846	CAGGAGCCCGGTGCTATGCCCCAGGGGCAAGCTGTGCGAGTCTTTCGGGCGAGCTGTGT	3905	3905
3835	CGGGGGTCTCCGTGTGTGTGCGCCCGGGGGCAGCCGCTGGGCATATTTAGGGCCGGGTGT	3894	3894
3906	GCTCTCGGGGGTGGCTAAGTCCATAGATTTTATCCCGCTTGAGACACTTCGACATCTGTC	3965	3965
3895	GCA CCGGTGGAGTGGCTAAGGGGTGGACTTTATCCCTGTGGAGAACCTTAGAGACAACA	3954	3954
3966	CGCGTCCCCCACTTTTAGTGACAAACAGCACCAACCTGTGTGTGCCAGACCTATCAGG	4025	4025
3955	TGAGTCCCCGGTGTTCACGGATAACTCTCTCTCCACAGTAGTAGTGTCCCGAGAGCTTCC	4014	4014
4026	TCGGGTACTTCGATGCCCGGACTGGCAGTGAAGAGAGACCAAGTTTCTGTGCGATATG	4085	4085
4015	TGGTCTACCTCCATGCTCTCCACAGGACGGCGCAAAAGCACCAAGTCTCCGGCTGCATATG	4074	4074
4086	CTGTCTAGGGGTATAAAGTGTAGTGTCTTAATCTCTCAGTGGCTGCCACCTCGGGGTGTG	4145	4145
4075	CAGTCTAGGGGTATAAAGTGTAGTACTCAACCTCTGTGTGCTGCAACACTGGGCTTTG	4134	4134
4146	GGCGTACTTGTCTAAAGGCAATGGCATCAATTCACCAATTAGGACTGGAGTCAAGACTG	4205	4205
4135	GTGCTTACATGTCTCAAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGAACAA	4194	4194
4206	TGACGACGGGGCGCCCATCAGTACTTCCACATATGGCAAAATCTCTCGCCGATGGGGCT	4265	4265
4195	TTACCATGGGACGCCCATCATGTACTTCCACTACGGCAAGTTCTGTCCGACGGCGGT	4254	4254
4266	GTGCGGGCGGCGCTTACGACATCATATGTGAATGCAATGCCCTGGGACTCTTACCA	4325	4325
4255	GCTCGGGGGCGCTTATGACATATATATTTGTGACGATGCGCATCTCCACCGATGCGCAT	4314	4314
4326	CCATCTTTGGCATCGGAACAGTCTCTTGATCAAGCAGAGACAGCTGGGGTTCAGACTG	4385	4385
4315	CCATCTTGGKCATCGGCATCTGCTCTTGACCAAGCAGAGACTCGGGGCGGAGACTGGTTG	4374	4374
4386	TGCTGGCTACAGCTACGGCCCTCTGGTTCAGTGTACAAACCCCAACATAGAGGAGG	4445	4445
4375	TGCTCGCACCGCACCCCTCGGGCTCCGTCACTGTGCGCCCATCTCCAACTCGAGGAGG	4434	4434

QY	4446	TGGCCCTTGGCAGGAGGCGAGATCCCTTCTATGGGAGGCGATTCCCTGCTTACA	4505	5526	CCGAGATGCTGAAGTCCAAGATCCAAAGGCTTATTGCGAGCAAGCTTCCAAACAAAGCTCAAG	5585
DB	4435	TTGCTCTGTGCCACACCGGAGAGATCCCTTTTACGGAGAGGCTATCCCTCTCGAAGTAA	4494	5515	CCGAGCAGTTCAAGCAGAGAGGCCCTCGGCTCTCTGAGACCGGCTCCCTGAGGAGAGG	5574
QY	4506	TCAAGGGAGGAGACATCTGATCTTCTGCGCATTTCAAGAAAGTGTGACGAGCTCGCGG	4565	5586	ACATCAACCCCATCTGTGAGGCTTATGGCCCAAGGTAGAACAAATTTCTGGGCCAAACACA	5645
DB	4495	TCAAGGGGGGAGACATCTCATCTTCTGTCTATTTCAAGAGAGTGTGCGAGACTCGCGG	4554	5575	TTATCGCCCTGCTGTCCAGACCAACTGGCAAAACCTCGAGACCTTCTGGGGGAGACATA	5634
QY	4566	CGGCCCTTCGGGGTATGGGCTTGAATCTAGTGGCATCTACAGAGGTTGGAGTCTCCG	4625	5646	TGTTGAACTTCAATTAAGCGCATCCATACCTCGCAGGATCTACACACTGCCAGGGAACC	5705
DB	4555	CAAAAGCTGTGCGCATTTGGGCATCAATGCGGTGGCTTACTACCGCGTCTTGAAGTCTCG	4614	5635	TGTGAACTTCAATCACTAGTGGGATACAAATCTTGGCGGCTTGTCAACGCTGCCTGTGTAACC	5694
QY	4626	TATACCAACTCAGGAGAGCTAGTGTGCTGCGCACCGACCGCTCATGACAGGGGTATA	4685	5706	CTGCACTAGCTTCCATGATGGGCTTCACTGCGCGCCCTCAACAGTCCGCTGTCAACAGCA	5765
DB	4615	TCATCCGACCAAGCGGCGATGTTGTGCTGTGGCAACCGATGCGCTCATGACCGGTATA	4674	5695	CGGCCATTCCTTCAATTTGATGGCTTTTACAGCTGCTGTCAACAGCCACTAAACACTAGCC	5754
QY	4686	CTGGGACTTTGACTCCGCTGATCGACTGCAACGTAGCGGTCACTCAAGTTGTAGACTTCA	4745	5766	CCACTATCTTCTCAAACTTTTGGGGGCTGTGCTAGCATCCCAAAATTTGACACCCCGCGG	5825
DB	4675	CGGGCACTTCGACTCGGTGATAGACTRCAATACGTGTGTCAACCCAGACAGTCGATTTCA	4734	5755	AAACCTCTCTTCAACATATTTGGGGGGTGGGTGGCTGCCAGCTCGCGCCCGCGGTG	5814
QY	4746	GTTTACACCCCAATTCACCAATTAACACAGATTTGCTCCTCAAGAGCTGTCTACGTA	4805	5826	GGGCCACTGGCTTCTGTGTCAGTGGCTAGTGGAGCTGCCGTAGGACGTATAGGCTTAG	5885
DB	4735	GCCTTGACCTTACCTTTCASCAITTGAGACAATCAAGCTTCCCGAGGATGTGTCTCCGCA	4794	5815	CCGCTACTGCTTGTGGCGCTTGGCTTAGCTGGCGCCCATCGGAGTGTGGACTGG	5874
QY	4806	GCCAGCGCGGGTCCGAGGTAGGGAGACTGGGAGACTGGGCAATTTATAGTATGTTTCCACTG	4865	5886	GTAAGGTGTAGTGGACATCTCTGGCAGGCTATGTTGCGGGCAATTTTCGGGGGCTCTCGTCG	5945
DB	4795	CTCAACGTGTGGGCGAGACTGGCAGGGGAGCCAGGCATCTACAGATTTGTGGCACCGG	4854	5875	GGAGGTCTCTATAGACATCTTTGAGGGTATGGCGCGGGCTGGCGGAGCTCTTGTGG	5934
QY	4866	GTGAGCGAGCTCAGGAATGTTTTCACAGTGTAGTGTCTGTGAGTGTCTACGACGAGGGG	4925	5946	CAATCAAGATCATGTCTGGCGAGAGCCCTCATATGAGAGATGTCTGTCATCTTCTGCTCG	6005
DB	4855	GGGAGCGCCCTTCGGCATGTTTGCATCGTCCCTCTCTGTGAGTGTATGACGAGGCT	4914	5935	CAATCAAGATCATGAGCGGTGAGTCTCCCTTCCACGCGGAGCTTGTGTCATCTACTGCCCG	5994
QY	4926	CCGATGTTATGAGTCAACCATCGAGACCAACCGTCAAGCTCAGGCTCAGGGGTATTTCAACA	4985	6006	GAATCTGTCTCGGGTGTCTTGTGAGTGGAGTCACTCGCGCGCATTTCTGCGCGGAC	6065
DB	4915	GTGCTTGGTATGAGTCAACCGCGCGAGACTACAGTTAGGCTTACGAGCGTACATGAACA	4974	5995	CCATCTCTCGCCCGGAGCCCTCGTAGTGGCGGTGTCTGTGAGCAATCTACTGCGCGGC	6054
QY	4986	CGCCCGTGTGCTGTGTCAGAGACCATCTTCAGTTTGGGAGGAGTCTTTCACCGGCC	5045	6066	AGTGGGACCGGGGAGGCGCGCTCCAAATGGATGAATAGACTCATTTGCTTTGCTTCCA	6125
DB	4975	CCCCGGGCTTCCCGTGTGCGAGACCATTTTGAATTTTGGGAGGGGCTTTTACAGGCC	5034	6055	ACGTTGGCCCGGCGGAGGGGAGTGCAGTGGATGAACCGGCTGATAGCTTCGCTCCC	6114
QY	5046	TCACACATAGATGCCACTTCTTTCCCAACAAAGCAATCGGGGAAATTTTCGCTAT	5105	6126	GAGGAAATCACTGTCGCCCCCAACCTACGACGAGTGGATGGATGGCTGCGAGCTGTGA	6185
DB	5035	TCATCATATAGATGCCACTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTCT	5094	6115	GGGGGAAACCATGTTTCCCCCAGCACTAGCTGTGCGGAGAGCGATGACGCTGCCCGGTCA	6174
QY	5106	ACTTAAACAGCTTACAGGCTTACAGTGTGCGCTAGGGCCAAAGCCCCCCCCCTCTGGG	5165	6186	CCCACTACTTGGCTTCTTACCAATACCAAGCTGTCTCAGAGACTTCCCAACTGGATTA	6245
DB	5095	ACCTGGTAGCTTACCAAGCCACCGTGTGCGCTAGGGCTCAAGCCCTTCCCCCATCGTGG	5154	6175	CTGCCACTCTGCGTCCCTTACCAACCCAGCTCTCTGAGGCGGACTGACACAGTGGATAA	6234
QY	5166	ACGTCTAGTGGAGTGTGACTCGACTCAAGCCACACTCGTGGGCCCCACACCTCTCC	5225	6246	CTGAGGACTCGCCCATCCCATCGCGGCTCGTGGCTCGCGCATGTGTGGGACTGGGTTT	6305
DB	5155	ACCAGATGTGGAGTGTGATGTCGCTCAAGCCCACTTCCATGGGCGCAACCCCTGCG	5214	6235	GCTCGGAGTGTACCACTTCCATGCTCGGTTCTTGGCTAAGGGACATCTGGGACTGGATAT	6294
QY	5226	TGTACCGCTTGGGCTGTGTTACCAACAGAGTCAACCTTCAACATCCCGTGTGACAAATACA	5285	6306	GCACCATCTTAAACAGCTTTTAAATTTGGTGTGACTTCCAAATTTATTTCCCAAGATGCCG	6365
DB	5215	TATACAGACTGGGGCTGTTTCAAGTGAATCACCTGTGACGACCCAGTCAACCAATACA	5274	6295	GCAGGTGTGTGAGCGACTTTTAAAGCTTAAAGCTTAAAGCTCATGCCACAGCTGCCCTG	6354
QY	5286	TGCGCACTGTGCAAGCCGACTTGTGAGTGTATGACGACCATATGCGTCTTGGCAGGGG	5345	6366	GCTTCCCTTGTCTCTGTCAAAAGGGGTACAAGGGCGTGTGGCGCGGCACTTGGCATCA	6425
DB	5275	TCATGACATGATGTGCGGCGGACTGTGAGGTCTGTACGAGCACTTGGGTGCTCGTTGGCG	5334	6355	GGATCCCTTGTGTCTTCCAGCGCGGTATTAAGGGGTCTGGCGAGGAGCGGATCA	6414
QY	5346	GAGTCTTGGCGCGCTCGCGCTTATGCTGCGGACCGGCTGTGTTGATCATCTCGGCC	5405	6426	TGACCAACAGGTTCTTGGCGGCAATATCTCTGGCAATGTCCGCTTGGGCTTCCATGA	6485
DB	5335	CGGTCTTGGCTGCTTTTGGCCGCGTATTTGCTGTCAACAGGCTGCGTGGTCTATGATGGCA	5394	6415	TGCACACTCGTGCACCTGTGGAGCTGAGATCACTGGACATGTCAAAAACCGGACGATGA	6474
QY	5406	GCTTGCACATTAACACGAGCGCTGCTTGGCGGAGCAAGAGGCTTCTTATGAGCTT	5465	6486	GAATCAAGGGGCTTAAGACCTTGAATATCTGGCAGGGGACTTTCCTTATCAATTTGT	6545
DB	5395	GGGTCTTGTTCGGGAAGCCGCAATCATCTGACAGGAGTCTCTACCGAGAGT	5454	6475	GGATCTGCTCTTAGGACCTTGAGGAGCAATGTGGAGTGGGACTTTCCTCCCAITTAATGCTC	6534
QY	5466	TTGATGATGAGGAGGATGTGCTCTTAGGGGCGCTCTCATTTGAGAGGGGCGAGGATAG	5525	6546	ACACGAGGCGGAGTGGGTGCGGAAACCCCGCGCAACTTTTAAGTTCGCTTCTGGAGGG	6605
DB	5455	TCGATGATGAGGAGAGTGTCTCTCAGCACTTACCGTACATCGAGCAAGGATGATGCTCG	5514	6535	ACACGAGGCGGCTTGTACCCCTTCTTGGCGCGGAACTACACGCTTTCGCGCTATGGAGG	6594
				6606	TGGCGGCTCAGAGTACGGGAGGTGACGAGCAGCGGCTCATACCACTACATAACAGGAC	6665

6595 TGTCTGAGAGGAATATGTGAGATAAGGAGGTCGGGAGCTTCCACTAGTCGACGGTA 6654
6666 TCACCACTGATTAAGTCCCTGCGCAACTACCTCTCCGAGTCTTTCTCTGGG 6725
6655 TGACTACTGAACTCTCAAAATGCGCGGTCAGAGTCCCATCGCCGAATTTTTCACAGAAT 6714
6726 TGGACGAGTGCAGATCCATAGTTTGCCTCCCAACCCAGAGCGGTTTTCCTCGGAGTGG 6785
6715 TGGACGGGTGCGCTACATAGTTTGGCGCCCTCTCAAGCCCTTGTCTCGGGAGGAGG 6774
6786 TCTCGTTCTGGTGGCTTAATTCATTTGTGTGCGGTCCAGCTTCTCTTGGACCCCTG 6845
6775 TATCATTTAGAGTAGGACTCCAGAAATACCGGTAGGTGCGCAATTAACCTTGGAGCCCG 6834
6846 AACCCGACACAGAGTATGATGTCATGCTTAACAGATCCATCTCATATCACCGCGGAGA 6905
6835 AACCGAGCTGGCGGTGACGCTTCAATGCTCACTGATCCCTCCCATATAACAGAGAGG 6894
6906 CTGACGCGCGGCTTTAGCGCGGGGTCAACCCCATCCGAGGCAAGCTCTCTCGCGAGCC 6965
6895 CGCGCGGCAAGGTTGCGGAGGGATCAACCCCTCTGTGGCAGCTCTCTCGGTAGCC 6954
6966 AGCTATCGGCAACCTCTCGGAGCCACTTGCACCCACCGCAAGCTTATGATGTGG 7025
6955 AGCTATCGGCTCCATCTCTCAAGCAACTTGCACCGCTAACCATGACTCTCCCTGATGCTG 7014
7026 ACATGTTGGATGCTAACT-----GTTTCATGGGGGGGATGATGCTCGATAG 7073
7015 AGCTCATAGAGGCGCACTCTTATGGAGGAGGATGGCGGCAACATCACAGAGGTTG 7074
7074 AGTCTGGGTCCAAAGTGGTCTCTGACCTCTCGACCAATGTTGCGAAGAGAGCG 7133
7075 AGTCAGAAACAAAGTGTGATCTGACCTCTCTGATCCGCTTGTGGCGAGGAGAGG 7134
7134 ACCTTGAGCTTTCGATACCATCAAGATPAATGCTTCCCAAGAGAGGTTCCCAACCACTT 7193
7135 AGCGGAGATCTCGGTACCGCAGAAATCTCGGAGAGTCTCGGAGATTCCGCCAGCGCC 7194
7194 TAGCGCTGGGACCGCTGATTAACCAACCGCTTGGATGCGTGGAAAGGCCAG 7253
7195 TGCGGCTTGGCGCGCGGACTATTAACCCCGCTAGTGGAGAGCGTGGAAAGCCCG 7254
7254 ATTACCAACCGGCACTTGTGGGCTGTGCTCTCCCTCTCTAGGAAACCCCGAGC 7313
7255 ACTAGAACCACTGTGCTCATGGCTGTGCTTCCCTTCCACTTCAAGTCCCTCTCTGTC 7314
7314 CTCCTCCAGAGGCGCGGACAGTGGGCTTAAGTAGAGACTCCATAGAGATGCCCCCTTC 7373
7315 CTCGCTCTCGAAGAGCGAGCGTGTCTCACTGAATCAACCTTACTTACTTGCCTTGG 7374
7374 AACAGCTGGCCATTAAGTCTTTGGCGAGCGCCCGGAGGATTCAGGCTTTTCCA 7433
7375 CCGAGCTCGCCASAGAAAGCTTTGGAGCTTCCACTTCCGCAATTAACGGGCGCAATA 7434
7434 CGGGGCGGCGCTGCGGATTTCCGCGAGTCAGAGCGCTCTGATGATGCTGGCCCTTTCGG 7493
7435 CGAACATCTCTG-----AGCGCGCCCTTCTGGCTGCCCCCGACTTCG 7482
7494 AGACAGGTTCCATCTCTTCCATGCCCCCTCGAGGGGAGCTTGGAGATCCAGAGCTGG 7553
7483 AGCTGAGTCTTWTKCTCATGCCCCCTGGAGGGGAGCTTGGGATCCGGATCTTA 7542
7554 AGCTGAGCAGGTAGAGCCCAACCCCGGCGGGGGTGGAGCTTCCCGGCTCGG 7613
7543 -----GCGAGC 7548
7614 ACTCGGCTCTGCTACTTCTCGAGGAGGAGCTCCGCTGCTGCTGCTCCATGT 7673
7549 GGTATGCTCAACGGTCAGTAGTAGGCGCAACCGGAGGATGTGCTGCTGCTCAATGT 7608
7674 CATACTCTGGACCGGCTCTAATAACTCTTGTAGTCCCGAAGAGGAGGATTAACCGA 7733

7609 CTTTACTTTGGAGAGCGCACTCGTCAACCCGCTGGCGCGGAGAACAGAAACTGCCCCA 7668
7734 TTAACCCCTTTGAGCACTCCCTGTTGCGATATACAAAGAGTGTACTGTACCAACAA 7793
7669 TCAATGCACTAAGCACTGTTGCTACGTACCAACAAATTTGGTGTATTCACCACTCAC 7728
7794 AGAGCGCTCTCACTAAGGGCTAAAAAGGTAACTTTTGTATAGGATGCAAGTCTCGACTCT 7853
7729 CGAGTGTGTTGCCAAAGGAGAAAGTCACTTTTGACAGACTGCAAGTCTTGACAGCGC 7788
7854 ACTACGACTCACTTTAAGGACATTAAGCTAGCGGCTCCAGAGTCAACCGAGGCTCC 7913
7789 ATTACAGGAGCTTCAAGGAGGTTAAAGAGCGGCGTCAAAAGTGAAGGCTTAACCTTBC 7848
7914 TCACCATGAGAGGCTTGCAGTTAAACCCCAACCCCATTTCTGACAGATCAAAATATGGGT 7973
7849 TATCGTAGAGAGCTTGCAGCTGACGCCCCCACTACAGCAATCCAGTTTGGTT 7908
7974 TTTGGGCTTAAGAGGCTCCGAGCTTTGTCGGGAGGCGCTTAAACCAATCAAGTCCGCTGT 8033
7909 ATGGGCAAAAGAGCTCGTTCGCTAGGAGGCGCTTAACCCACATCACTCCGCTGT 7968
8034 GGAAGAGCTCTCTGAGGAGCTCAGAAACACCAATTTCCCAACCAATTTGGCCAAATG 8093
7969 GGAAGAGCTTCTGGAAGCAATGTAAACCAATAGACACTACCATCATGGCTTAAGAACG 8028
8094 AGTGTCTTCGCTGAGCCCAACAGGGGGGCAAGAAACAGCTCCCTTATGTTTACC 8153
8029 AGTGTCTTCGCTGAGCTGAGAGGGGGTGTAAAGCAGCTCTCTCATCTGTTTCC 8088
8154 CTGACCTCGGCTCAGGGTCTCGAGAGATGGCCCTTTATGACATTACACAAAACCTTC 8213
8089 CCGATCTGGGCTGCGGTGCGAAGATGGCTTTGACGAGTGTGTACAAAGCTTC 8148
8214 CTCAGCGGTGATGGGGGCTTCTTATGATTCAGATTTCCCGCTCAGCGGTTAGAGT 8273
8149 CTTGGGCTGATGGGAGCTCTTACGATTTCAATATCTACAGAGAGCGGTTGAAT 8208
8274 TTTCTTGAAGCATGGGCGGAAAGAGGACCTATGGTGTTCGTATGATACCCGAT 8333
8209 TCTCTGTGAGCGTGGAGTCAAGAAACCCAAATGGGTTCTGATGATACCCGCT 8268
8334 GCTTTGACTCAACCGTCACTGAGAGAGCATCAGGACTGAGGAGTCCATATATCGGCT 8393
8269 GCTTTGACTCCACAGTCACTGAGAGCGCATCGTACGAGGAGGCAATCTACCAATGT 8328
8394 GCTCTTGGCGAGGAGGCGCAGCTGCGCATACATGCTGTACTGAGAGACTTTAGTGG 8453
8329 GTGACCTCGACCCCAAGCCCGGTGGCCATCAAGTCCCTCACCGAGAGGCTTTATGTTG 8388
8454 GAGGCTTATGTTCAACAGAGGCGCAACCTGCGGGTACAGGCTTCCCGCCAGCG 8513
8389 GGGGCTTCTTACCAATTCARGGGGGAGACTGCGGCTTATCGAGGTCCCGCGAGCG 8448
8514 GGTGTCTCACTAGCTAGCTAGGAGGAGCAACCATCACATGCTTACGTGAAGCTTTAGCGGCTT 8573
8449 GGTGTCTGACACTAGCTGTTGTGATACCCCTCACTTGTACATCAAGGCGCGGAGCT 8508
8574 GTAAAGCTGAGGATTAATTCGCGCCCAATGCTGTGTATGCGGCGATGATCTTGTGTTCA 8633
8509 GTGAGCGCGAGGCTTCCAGGACTGCAACATGCTGTGTGTGCGGAGCTTTAGTCTTTA 8568
8634 TCTCAAGAGCGGAGGAGCGGAGGAGCGGAGGAGCTGAGAGCTTTCAGGAGGCTA 8693
8569 TCTGTGAAGCGCGGGGTTCCAGGAGGAGCGCGGAGCTTTCAGGAGGCTA 8628
8694 TGACAGGTTATTTGCTCTCTCTGTTGACCCCGCCAGAGCGGAGTATGATCTGAGGCTGA 8753
8629 TGACAGGTTATCTCGCGCCCCCTTGGGAGCCCCCAACACAGATACAGTCTTGGAGCTCA 8688
8754 TAAATCTTGTCTCTCAAAATGTTGTGTTGGCGCTGGGCGCAACAGGCGCGCGAGATCT 8813
8689 TAAATCTATGCTCTTCAACGTTGTCAAGTGTGAGTGTGCGGCTGGAAGAGGCTTACT 8748

Query Match	45.6%	Score 4425.8;	DB 1;	Length 9401;
Best Local Similarity	67.3%;	Pred. No. 0;		
Matches 6364;	Conservative 0;	Mismatches 2997;	Indels 90;	Gaps 5;

QY	6	CCCTAATAGGGCGCACACTCCGCCATGAATCACTCCCTGTGAGGAAC	TACTGTCTTCA	65
Db	7	CCCTGATGGGCGCACACTCCACCATGAATCACTCCCTGTGAGGAAC	TACTGTCTTCA	66
QY	66	CGCAGAAAGCGTTACGCATGGCGTTAGTATGATGTCTGACACCTC	CAGGCCCCCCCCC	125
Db	67	CGCAGAAAGCGTTACGCATGGCGTTAGTATGATGTCTGACACCTC	CAGGACCCCCCCC	126

Copied from 09540843 on 05/19/2004

Qy	3426	TTCTCGCCCCCATCACTGCTTAGCGCCAGCAGACAGCTGGCCCTTTTGGGCACCATAGTGG	3485	Qy	4506	TCAGGGAGGAGAGACATCTGATCTTCTGCAATTCAGAGAAAAGTGTGACGAGCTCGCGG	4565
Db	3415	TGCTGGCGCCATCATCGGGGTACGCGCAGCAGACAGAGGGGCTCTTAGGGTGATATCA	3474	Db	4495	TCAGGGGGGAGACATCTCATCTTCTGCAATTCAGAGAGAGTGCAGAGAACTCGCGG	4554
Qy	3486	TGAGCATGACGGGGGGGACAGACAGAGACAGAGCTGGGAAATTCAGGTCTGTGCACAG	3545	Qy	4566	CGGCCCTTCGGGGTATGGGCTTGAACTCAGTGGGCATCTACAGAGGTTTGGAGCTGCGG	4625
Db	3475	CCAGCCTTAACGTGGCCGGGACAAAACCAAGTGGGGGTGAGGTCCAGATTTGTCAACTG	3534	Db	4555	CAAAAGCTGTGCGCATTTGGGCATCAATGCGGTGGCCCTACTACCGCGGTCTTGACGCTGCG	4614
Qy	3546	TCACTCAGTCTTCTCGGAACATCCATCTCGGGGGTTTGTGGACTGTCTACCATGGAG	3605	Qy	4626	TAAATACCAACTCAGGGAGACGTAGTGGTGGCCGACCGACCGCCCTCATGACAGGGTATA	4685
Db	3535	CTGCGCAACCTTCTTGGGCACTGTGATCAATGGGGTGTCTGGACTGTCTACCGGGG	3594	Db	4615	TCATCCGACGACGGCGCATTTGTGCTGGGCAACGATGCCCTCATGACCGGCTATA	4674
Qy	3606	CTGGCAACAGACTCTGGCGGCTCACGGGGTCCGGTCAAGCAGATGTACTCCAGTGTCTG	3665	Qy	4686	CTGGGGACTTTGACTCCGATTCGATCGACTGCAACGTAGCGGTCACTCAAGTTGTAGACTTCA	4745
Db	3595	CCGGAACGAGGACCATCGGCTCACCAAGGGTCTCTGTCATCCAGATGTATCAATGTAG	3654	Db	4675	CGCGGACTTTCGACTTCGGTGTAGACTGCAATACGTTGTCAACCCAGACAGTTCGATTTCA	4734
Qy	3666	AGGGGACTTAGTAGGGTGGCCAGCCCTCGGACTTAAATCTTTTGGAGCGGTGCACGT	3725	Qy	4746	GTTTAGACCCCACTTCAACCATACACACAGATTTGTCTCAAGACGTGTCTCACGTA	4805
Db	3655	ACCAAGACCTTTGGGCTGGCCCGCTCCGCAAGGTAGCCGCTCAITTGACACCCCTGCATTT	3714	Db	4735	GCTTTGACCCCTACCTTTCAACATGAGACAATCAACGCTCCCCAGGATGCTGTCTCCGCA	4794
Qy	3726	GTGGAGCGGTGACCTGTACTCTGTCACGCGGAACGCTGATGTCAATCCCGGCTCGAAGAC	3785	Qy	4806	GCCAGCGCGGGTCCGACCGGTAGGGGAAGACTGGGCAATTTATAGGTATGTTTCCACTG	4865
Db	3715	GCGCTCTCGGACCTTTTACCTGGTCAGAGGCACCGCATGTCTTCCCGTGGCGCGG	3774	Db	4795	CTCAACGTGGGGCAGGACTGGCAGGGGAAGCCAGGCATCTACAGATTTTGGGACCCGG	4854
Qy	3786	GCGGGGACAAACGGGGAGCGCTACTCTCCCGGAGACCTCTTTCCACCTTTGAAGGGTCTCT	3845	Qy	4866	GTGACGAGCCCTCAGGAATGTTTGAAGTGTAGTGTCTCTGTGAGTGTCTACGACGAGGGG	4925
Db	3775	GGGGTGATAGCAGGGGCGCTGCTGTGCGCCCGGCCATTTCTTACTTGTAAAGGCTCTCT	3834	Db	4855	GGGAGGGCCCTCCGCGCATTTGACTCGTCCGTCCTCTGTGAGTGTCTATGACGAGGCT	4914
Qy	3846	CAGGAGCCCGGTGTATGCCCCAGGGGCGACGCTGTGCGAGTCTTCCGGGCGAGTGTGT	3905	Qy	4926	CGGCATGTATGAGTGTCAACCATCGGAGACACCGTCAAGGCTCAGGGGCTATTTCAACA	4985
Db	3835	CGGGGGTCCGCTGTTGTGCCCCCGGGCAGCCGCTGGGCATATTTAGGGCGCGGTGT	3894	Db	4915	GTGCTTGGTATGAGTGTCAACCGCGGAGACTACAGTTAGGCTACGAGCTACATGMAACA	4974
Qy	3906	GCTCTGGGCGGTGAGTCAATAGATTTCATCCCGTGGAGACACTCGACATCGTCA	3965	Qy	4986	CGCCCGGTTTGGCTGTGGCCAGACCATCTGAGTGTGGGAGGAGTGTTCACGGCC	5045
Db	3895	GCACCGGTGGAGTGGTAAAGCGGTGAGCTTTATCCCTGTGGAGAACCTTAGAGACAACA	3954	Db	4975	CCCCGGGGCTTCCCGTGTGCCAGGACCATCTTGAATTTTGGGAGGGCTTTTACAGGCC	5034
Qy	3966	CGCGGTCCCCACCTTTTAGTGACACAGCAGACACCTGCTGTGCCCCAGACCTATFCAGG	4025	Qy	5046	TCACACATAGATGCCACCTTCTTCCAAAACAAAGCAATTCGGGGGAAATTTTCGCAT	5105
Db	3955	TGAGTGTCCCGGTGTTACGGATATCTCTCTCCACGATGATGTCGCCAGAGTTCAGG	4014	Db	5035	TCACTCATATAGATGCCACCTTCTATCCAGACAAGAGAGTGGGAGAACCTTCTCTT	5094
Qy	4026	TGCGGTACTTTCATGCCCCGACTGGCAGTGGAAAGAGCACAAAGTTCTGTGCGCATATG	4085	Qy	5106	ACTTAAACGCTTACAGGGTACAGTGTGCGGTAGGGGCAAGCCGCCCTCCCGCTGTGGG	5165
Db	4015	TGGCTCACCTTCCATGCTCCACAGGCGGGCAAAAGCACCAAGGTCCCGGCTGCATATG	4074	Db	5095	ACCTGTAGCGTACCAAGCCACCGTGTGCGGTAGGGCTCAAGCCCTCTCCCATCTGGG	5154
Qy	4086	CTGCTCAGGGGTATAAGTGTAGTGTAACTTAACTCCCTCAGTGGCTGCACCTGGGGTTG	4145	Qy	5166	AGTCTCATGTGGAAGTGTGTGACTGACCTCAAGCCACACACTGTGGGGCCCAACCTCTCC	5225
Db	4075	CAGCTCAGGGCTATAAGGTGTAGTACTCAACCCCTCTGTGTGCAACACTGGGCTTTG	4134	Db	5155	ACCAGTGTGAAGTGTGTGATTCGCTCAAGCCCAACCTCCATGGGCAACACCCCTGCG	5214
Qy	4146	GGCGGTACTTGTCTAAGGCACATGGCATCAATCCCAACATTAGGACTGGAGTCAGGACTG	4205	Qy	5226	TGTACCGTGGGCTCTGTATACACAGAGTCAACCTCACACATCCCGTGACGAATACA	5285
Db	4135	GTGCTTACATGTCCAAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGNACA	4194	Db	5215	TATACAGACTGGGCGCTGTCAGAATGAATCAACCTGACGACCACTACCAATACA	5274
Qy	4206	TGACGACCGGGGCGCCATCACTACTCAACATATGGCAAAATCTCGCCGATGGGGCT	4265	Qy	5286	TGCGCACTGTATGACGACCGACCTTGAAGTCAAGACGACATGAGTCTTGGGAGGGG	5345
Db	4195	TTACCACCTGGCAGCCCATCACGTACTCCACCTACGCGAAGTTCTTTCGCGAGCGGGT	4254	Db	5275	TCATGACATGATGTGGCGGACCTTGAAGTGTGCTCACGAGCACTGGGTGCTGCTGGCG	5334
Qy	4266	GTGCGGGCGGCTACGACATCATATGTGATGAATGCCATGCCGTGGACTTACCA	4325	Qy	5346	GAGTCTTGGCGCGCTCGCGGGTATTCCTGGGCAACCGGGTGTGTTGTCATCATCGGCC	5405
Db	4255	GCTCGGGGGCGCTTATGACATATAATTTGTGACAGGTGCCACTCCACGGATGCCACAT	4314	Db	5335	CGCTCTGGCTGTGCGCGGTATTCCTGTCAACAGGCTGTGCTCATAGTGGGCA	5394
Qy	4326	CCATCTTGGCATCGGAACAGTCTTGTCAAGCAGAGACAGTGGGGTCAAGTAACTG	4385	Qy	5406	GCTTGCAATTAACGAGCGAGCGGTGCTGGCGGACAGAGGTCTCTATAGGGCTT	5465
Db	4315	CCATCTTGGCATCGGCATGTCTTGAACAGCAGAGACTGGGGGGGAGACTGGTTG	4374	Db	5395	GGGTGCTTGTTCGGGAAGCGGGCAATCATCTGACAGGGAGTCTCTTACCGAGAGT	5454
Qy	4386	TGCTGCTACAGTACGCCCCCTGGGTGAGTGACAAACCCCAACCCCAACATAGAGGAGG	4445	Qy	5466	TTGATGAGATGGAGGAATGTGCTCTTAGGGCGGTCTCATTTGAAGAGGGCAGCGGATAG	5525
Db	4375	TGCTCGCCACCGCCACCCCTCGGGCTCCGTCACTGTGCCCCATCCCAACATCGAGGAGG	4434	Db	5455	TGATGAGATGGAAGAGTGTCTCAGCACTTACCGTACATCGAGCAAGGATGATGCTCG	5514
Qy	4446	TGGCCCTTGGGAGGAGGGGAGATCCCTTCTATGGGAGGGGATTTCCCTGTCTTACA	4505	Qy	5526	CCGAGATCTGAGTCCAGATCCAAAGCTTTTGCAGCAAGCTTCCAAAACAGCTCAAG	5585
Db	4435	TTGCTCTGTCCACACCGAGAGATCCCTTTTACGGCAAGGCTATCCCGCTCGAAGTAA	4494	Db	5515	CCGAGGCTTCAAGCAGAAAGGCCCTCGGCCCTCTGACAGACCGCTCCCGTCAGGAGG	5574
				Qy	5586	ACATACAACCCACTGTGCAAGGCTTTCATGGGCCCAAGGTAGAACAAATTTCTGGGCCAAACACA	5645

7794 AGAGCGCTCACTAAGGCTTAAAGGTAACCTTTTGTAGATGCAAGTCTCGACTCTT 7853
 7729 GCAGTGTCTGCAAGGCAAGAGTACATTTGACAGACTGCAAGTCTTGACAGCC 7788
 7854 ACTACAGCTCAGTCTTAAAGGACATTAAGTAGCGGCTTCAAGGTCACCGAAGGCTCC 7913
 7789 ATTACAGGACGTACTCAAGGAGGTTAAAGCAGCGGCTCAAAAGTGAAGGCTTAACTTC 7848
 7914 TCACCATGAGGAGGCTTGCAGTTAAACCCCACTTCTGCAAGATCTTAATATGAGT 7973
 7849 TATCGTAGAGGAGCTTGCAGCTGACGCCCCACACTCAGCAAAATCCAAAGTTGGTT 7908
 7974 TTGGGGCTAAGAGGCTCCGAGCTTGTCCGGGAGGCGCTTAAACACATCAAGTCCGCTGT 8033
 7909 ATGGGCAAAAGACGCTCGTTCATCCAGAAAGCGCTTAAACACATCAAGTCCGCTGT 7968
 8034 GGAAGACCTCTGAGGAGTCAAGAACCAATTCACACACATTTAGGCGCAAAATG 8093
 7969 GGAAGACCTCTGAGGAGCAATGTAAACCAATAGACACTACCATCATGGCTTAAAGACG 8028
 8094 AGGTGTTCTGCGTGGACCCCAACAGGGGGCAAGAACGAGCTCCGCTTATCGTTTACC 8153
 8029 AGGTGTTCTGCGTTCAGCTGAGAGGGGGTCTGTAAGCAGCTCGTCTCATCGTGTTC 8088
 8154 CTGACCTCGGCGTCAAGGCTCTGCGAGAGATGGCCCTTTTATGACATTTACACAAAATCTT 8213
 8089 CCGATCTGGGCGTGGCGTGTGCGAAGAGATGCTTTGTACGAGCTGGTTACAAAGCTCC 8148
 8214 CTGAGCGGTGATGGGGCTTCTTATGATTCAGTATTCCTCCGCTCAGCGGGTAGAGT 8273
 8149 CTTTGGCGGTGATGGGAAAGCTCTACGATTCCTCAATTAATCAACGAGCAGCGGGTTGAAT 8208
 8274 TTTCTTGAAGCATGGCGGAAAGAACGACCTTATGGTGTGTTTGTATGATACCGAT 8333
 8209 TCCTGCTGAAGCGTGGAGTTCAGAAACCCCAATGGGGTCTCGTATGATACCGCT 8268
 8334 GCTTTGACTCAACCGTCACTGAGAGAGACATCAGGACTGAGGAGTCCATATATCGGGCTT 8393
 8269 GCTTTGACTCAACGCTGAGAGAGACATCCTGAGGAGGAGGCAATCTACCAATGTT 8328
 8394 GCTCTTGGCCGAGGAGGCCACACTGCCATACACTCGTAACTGAGAGACTTTACGTGG 8453
 8329 GTGACCTCGACCCCAAGCGCGTGGCCATCAAGTCCCTCACCAGAGGCTTTATGTTG 8388
 8454 GAGGCGCTATGTTCAACAGCAAGGCGCAACCTGCGGGTACAGCGTTCGCGCGCCAGCG 8513
 8389 GGGGCGCTCTTACCAATTTCAAGGGGGGAGAACTGCGGCTATCGCAGGTGCGCGCAGCG 8448
 8514 GGTGCTCACCCTAGCATGGGGAACACCACTACATGCTAGCTGAAAGCCTTACGCGCTT 8573
 8449 GGTGCTGACAACTAGTGTGGTAAACCTCTACTTGTACATCAAGGCGCGGAGCCTT 8508
 8574 GTAAAGTGCAGGATAATCGCGCCCAACATGCTGTATGCGCGATGACTTGGTTGTCA 8633
 8509 GTGAGCGCGGAGGCTCCAGGAGTGCACCATGCTGCTGTGTCGAGCACTTATGCTTGA 8568
 8634 TCTCAGAAACCGAGGAGCGAGGAGGAGCGGAGCCTGAGAGCTTCCAGGAGGCTA 8693
 8569 TCTGTGAAGCGCGGGGCTTCAGGAGGAGCGCGGAGCCTGAGAGCCTTCCAGGAGGCTA 8628
 8694 TGACAGGATTTCTGCGCCCTCTGCTGACCCCGCCAGACCGGAGTATGATCTGGAGCTGA 8753
 8629 TGACAGGATTTCTGCGCCCTCTGCGGAGCCCGCCCAACACAGATATGAGCTTGGAGCTCA 8688
 8754 TAAACATCTTCTCTCAAAATGTGTGTGCGGCTGGGCGCCCAAGGCGCGCGAGATACT 8813
 8689 TAAACATCTTCTCTCAACGTGTGAGTGCAGTGCAGGCGGCTGGAAAGAGGCTTACT 8748
 8814 ACCTGACGAGAGACCTTACCACTTCCATTCGCGCGGCTGCTGGGAAACAGTTAGACACT 8873
 8749 ACCTCACCGTACCTTACCAACCCCTCTCGGAGAGTGTGCTGGGAGAGCAGCAAGACACA 8808

8874 CCCCTGTCAATTCATGGCTGGGAAACATCATCATCAGTACGCCCCGACCATATATGGCTCGCA 8933
 8809 CTTCAAGTCAATTCCTGGCTAGGCAACAATAATCATGTTTGGCCCCACACTGTGGCGAGCA 8868
 8934 TGGTCTGTATGACACACTTCTTCTCATCTCTCATGCTCAAGACAGCTGGACCAAGACC 8993
 8869 TGATACTGATGACCACTTCTTCTTAGCGTCTTATAGCCAGGAGCAGCTTGAACAGGCC 8928
 8994 TCAACTTTGAGATGTACGAGCGGTGTACTCCGTGAGTCCCTTGGACCTCCAGCTATATA 9053
 8929 TCGATTGCGAGATCTACGGGGCTCTACTCATFAGAACCACTTGTACTCTCCANTCA 8988
 9054 TTGAAGGTTTACATGGCTTGAAGCTTTTCTCTGACACATACACTCCCCAGAACTGA 9113
 8989 TTTCAAGACTCCATGGCTCAGCGCATTTTCACTCCACAGTTACTCTCCAGGTGAATTA 9048
 9114 CAGGGTGGCTTACGCTCAGGAACTTTGGGGCGCCACCTTCAGAGGCTGGAAGAGCC 9173
 9049 ATAGGGTGGCGCATGCTTCAGAAAATTTGGGGTACCGCCCTTGGAGCTTGGAGACAC 9108
 9174 GGGCAGCTGCTCAGTCAAGGGCTCCCTCATCTCCGCTGGGGGAGAGCGGCTTGGGCTC 9233
 9109 GGGCGGAGGCTTCCGCTAGGCTTCTGGCCAGAGGAGGAGGCTGCCATATGTGGCA 9168
 9234 GATATCTTCAATTTGGGGTGAAGACCAAGCTCAAACTCACTCCATTTGCCGAAGCC 9293
 9169 AGTACTCTTCACTGGGCGTGAAGAACTCAAGCTCAAACTCACTCAATAGCGCGCTG 9228
 9294 GCTCTGTGATTTTCCAGCTGGTTTCAAGCTGGCGCGCGCGGCGGCGGAGCATTTATCACA 9353
 9229 GCGAGCTGAGCTTGTCCGCTGGTTTCAAGGCTGGCTTACAGCGGGGAGACATTTATCACA 9288
 9354 GGTGTGCGTGGCGCCGAGCCCGCTTATGCTCTTTTGGCTACTCTCTACTTTTGTAGGGG 9413
 9289 GGTGTCTCATCGCGCGCGCTGATCTGTTTTTGGCTACTCTCTCTCTCTCTCTCTCTCT 9448
 9414 TAGGCTTTTCTACTTCCCGCTCGGTAGAG 9444
 9349 TAGGCTATCTACT 9379

RESULT 14

US-08-440-519-9
 ; Sequence 9, Application US/08440519
 ; Patent No. 5712087
 ; GENERAL INFORMATION:
 ; APPLICANT: Houghton, Michael
 ; APPLICANT: Choo, Qui-Lim
 ; APPLICANT: Kuo, George
 ; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
 ; TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Chiron Corporation
 ; STREET: P.O. Box 8097 (Int. Prop. R-440)
 ; CITY: Emeryville
 ; STATE: CA
 ; COUNTRY: U.S.A.
 ; ZIP: 94662-8097
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/440,519
 ; FILING DATE: 12-MAY-1995
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/910,760
 ; FILING DATE: 07-JUL-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Blackburn Esq., Robert P.

```

/ REGISTRATION NUMBER: 30,447
/ REFERENCE/DOCKET NUMBER: 0101.002
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (510) 601-2702
/ TELEFAX: (510) 655-3542
/ INFORMATION FOR SEQ ID NO: 9:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 9401 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 342..9374
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 366
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Arg."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 372
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Thr."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 867
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Thr."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 1341
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Val."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 2148
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ile."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 2883
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Asn."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 3691
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ser."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 3690
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Thr."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4167
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Leu."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4323
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Val."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4701
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Tyr."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4752

```

```

/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ser."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 5970
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Gly."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 6183
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be His."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 6186
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Cys."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 6402
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Val."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7386
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ser."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7494
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Phe."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7497
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ala."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7845
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Phe."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 8409
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Gly."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 9102
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Gly."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 9327
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Pro."
/ US-08-440-519-9

```

```

Query Match      45.6%; Score 4425.8; DB 1; Length 9401;
Best Local Similarity 67.3%; Pred. No. 0;
Matches 6364; Conservative 0; Mismatches 2997; Indels 90; Gaps 5;

QY 6 CCCCTAATAGGGGCGACACTCCGCATGATCACTCCCTGTGAGGAAGTACTGTCTTCA 65
Db 7 CCCCTGATGGGGCGACACTCCACCATGATCACTCCCTGTGAGGAAGTACTGTCTTCA 66
QY 66 CGCAGAAAGCGTCTAGCCATGCGTGTAGTATGATGATCGTACAGCCTCCAGGCCGCCCCC 125
Db 67 CGCAGAAAGCGTCTAGCCATGCGTGTAGTATGATGATCGTACAGCCTCCAGGCCGCCCCC 126
QY 126 TCCCGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCGGGAGAC 185

```

Db 127 TCCGCGAGAGCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAGGACGAC 186
 Qy 186 TGCGTCTCTTTTGGGATAAACCCACTCTATGCGCGGCCATTTGGCGGTGCGCCCGCAAGA 245
 Db 187 CGGGTCTCTTTTGGGATCAACCCGCTCAATGCTTGGAGATTTGGGGTGGCCCGCAAGA 246
 Qy 246 CTGCTAGCGAGTAGCGTTGGGTTGGAAAGGCGCTTGTGGTACTGTGCTGATAGGGTGT 305
 Db 247 CTGCTAGCGAGTAGTGTGGGTCGCGAAAGGCGCTTGTGGTACTGTGCTGATAGGGTGT 306
 Qy 306 GCGAGTGGCCCGGGAGGTCTGCTAGACCGGTGACCATGACACAAATCTTAACTCTAA 365
 Db 307 GCGAGTGGCCCGGGAGGTCTGCTAGACCGGTGACCATGACACAAATCTTAACTCTAA 366
 Qy 366 GAAAAACCAAAAGAAACCAACCGCTGCGCCCAAGACGTTAAGTTTCCGGGGCGCGGC 425
 Db 367 AAAAAACAAACGTAACCAACCGCTGCGCCCAAGACGTTAAGTTTCCGGGGCGCGGC 426
 Qy 426 AGATCGTTGGCGAGTATCTTGTGCGCGCAAGGCGCCCGAGTTGGGTGTCGCGCGCA 485
 Db 427 AGATCGTTGGCGAGTATCTTGTGCGCGCAAGGCGCCCGAGTTGGGTGTCGCGCGCA 486
 Qy 486 CAAGGAAGACTTCGAGCGGTCCCGACCGAGTGAAGGCGCGACCATCCCTAAAGATC 545
 Db 487 CGAGAAAGACTTCGAGCGGTCCCGACCGAGTGAAGGCGCGACCATCCCTAAAGATC 546
 Qy 546 GCGGCTCCACTGCGCAATCTCGGGAAACCGAGTACCCCTGCGCCCTATACGGGAATG 605
 Db 547 GTGCGCGAGGAGGAGGACTTGGGCTCAGCCCGGCTACCTTGGGCCCTCTATGGCAATG 606
 Qy 606 AGGGACTTCGGCTGGCGAGTAGTCTCTGTGCGCCCGAGGTTCCTGCTCTTGGGGCC 665
 Db 607 AGGGCTTCGGGTCGGCGGAGTGTCTCTGTCTCCCGTGGCTCTCGGCCCTAGCTGGGGC 666
 Qy 666 CCAATGACCCCGGCTAGTTCGGGAAACCGTGGTAAAGTTCATGATACCTTAAGTGG 725
 Db 667 CCACAGACCCCGGCTAGTTCGGGAAACCGTGGTAAAGTTCATGATACCTTAAGTGG 726
 Qy 726 GCTTTGCGGACTCATGGGTGATATCCCTGCTGCGCGCCCGCTCGGGCGTGCCTCA 785
 Db 727 GCTTTGCGGACTCATGGGTGATATCCCTGCTGCGCGCCCGCTCGGGCGTGCCTCA 786
 Qy 786 GAGCTCTTCGCGCATGCGGTGAGAGTCTCTGAGGACCGGGTTAAATTTGCAACAGGGA 845
 Db 787 GGGCCCTTCGCGCATGCGGTGAGAGTCTCTGAGGACCGGGTTAAATTTGCAACAGGGA 846
 Qy 846 TACCGGTTGCTCTCTTTTCTATCTTCTGTCGCGCTGCTGCTGCTGCTGCTGCTGCTG 905
 Db 847 TTCTGCTGCT 906
 Qy 906 TCTCCGCTGCGGAGTGAAGAACATCAGTACCGGCTACATGGTGAATGATGATGATGAT 965
 Db 907 CTTCCGCTTACCAAGTGCAGAACTTCAACGCGGGCTTTACACGCTGACCAATGATG 966
 Qy 966 ATGACAGCATTAATCTGCGAGCTCTCAGGCTCTCTCTCTCAGCTGCTGCTGCTGCTG 1025
 Db 967 ACTCGAGTATGTTAGAGCGCGGCGATGCCATCTGTCACACTCTCGGGGTGCTCTCT 1026
 Qy 1026 GCGAGAAAGTGGGAATGCAATCTCAGTGTGGATACCGGCTCTCAGCAATGTCGCGTGC 1085
 Db 1027 GCGTTCGTGAGGGCAACCGCTCGAGGTGTTGGGTGCGATGACCCCTACGCTGCGG 1086
 Qy 1086 AGCGGCGCGGCGCTCAGCAGGGCTTGGGAGCGACATCGACATGTTGATGATGATG 1145
 Db 1087 GGGATGGCAAACTCCCGCGGACGAGCTTCGACGTCACTGATGCTGTTGTCGGAGCG 1146
 Qy 1146 CCAGCTCTGCTCTGCGCTCTAGTGGGAGCTCTGCGGTGGGGTGAATGCTCGCAGCCC 1205
 Db 1147 CCACCTCTGTTGCGGCCCTCTAGTGGGAGCTATGCGGGTCTGCTCTTCTGTCGGCC 1206
 Qy 1206 AATGTTCTGTCGCGCGACGACCACTGTTGTTGTCGAAGTCTGCAATGTCCTCATCT 1265
 Db 1207 AACTGTTCACTTCTCTCCAGGCGCACTGGAGCGCAAGGTTGCAATGTTGCTCTCTCT 1266

Qy 1266 ACCCTGGTACCATCACTGGACACCGCATGGGATGGGACATGATGATGATGATGATG 1325
 Db 1267 ATCCCGCCCATATAACGGGTCAACCGATGGGATGGGATGATGATGATGATGATG 1326
 Qy 1326 CGGTACCAATGATCTTTGGGATAGGATGCGTGTCCCGAGGTCTATTATAGACATCAT 1385
 Db 1327 CGACGGCGTGTGTAATGCTAGCTGCTCCGATCCCAAGCCATCTTTGGACATGATCG 1386
 Qy 1386 GCGGGGCTCATTTGGGGCGTCAATGCTGCGCTTCTCTATGACAGGAGGTGGG 1445
 Db 1387 CTGGTCTCACTGGGAGTCTTGGCGGCTATAGCTATTCTTCCATGTTGGGAACTGGG 1446
 Qy 1446 CGAAAGTCTGTTGCTATCTTCTTGGCGCGCGGTGACCGCGGACCCCATCTACTGTTG 1505
 Db 1447 CGAAGTCTGTTGCTATCTTCTTGGCGCGGTGACCGCGGACCCCATCTACTGTTG 1506
 Qy 1506 GGGGTTTGGCGCGGACGACCGGGGCTTCAACAGCTTATTTGACATGCGGCGCCAGGC 1565
 Db 1507 GGGGAAGTGGCGGCGACACTGTGCTGGAATTTGTTAGCTTCTCGCACCGGCGCCAGC 1566
 Qy 1566 AGAAATCCAGTCTGTTTACACCAATGGGAGTGGGACATCAACCGCACCGGCTGAACT 1625
 Db 1567 AGAAGTCTGAGTGTATCAACCAAGCGGAGTGGGACCTCAATAGCACGGGCTGAACT 1626
 Qy 1626 GCAATGACTCTTGTGACACCGGCTTATGCGCTCTCTGTTCTACACCCACAGCTTCAA 1685
 Db 1627 GCAATGATAGCTTCAACCGGCTGTTGGGAGGCTTTTCTTCAACCAAGTTCAACT 1686
 Qy 1686 GCTCAGGATGTCCTCCGAAAGCATGTCGCTGCGCGGATGAGGCTTTCGCGGTGGGAT 1745
 Db 1687 CTTGAGGCTGCTCGAGAGGCTAGCGAGTGGCGGACCTTACCGATTTTGGACGAGCT 1746
 Qy 1746 GCGGCGCTTGGCAATATGAGGATATGTCACCAATCCAGAGGATATGACACCTATTGCT 1805
 Db 1747 GGGGCGCTTATGATGTTGCAACCGGAGCGGC-----CCCGACGAGCGGCTTACTGCT 1800
 Qy 1806 GGCACCTACCCACCAAGGAGTGTGGGTGCTCGCGGAAGACTGTGTGTGGGCCAGTGT 1865
 Db 1801 GGCACCTACCCACCAAGGAGTGTGGGTGCTCGCGGAAGACTGTGTGTGGGCCAGTGT 1860
 Qy 1866 ACTGTTTCAACCCCGGAGTGTGGGACGACGAGGCTTGGAGGCGCCACAT 1925
 Db 1861 ATTGCTTCACTCCGAGCGGCTGGTGGGAAACGACGAGGTCGGGCGGCGCCACCT 1920
 Qy 1926 ACAGTGGGGGAGATGAGACAGATGCTTCTTATTTGAACAGACTCCAGCACCGCTGG 1985
 Db 1921 ACAGTGGGGGAGATGAGACAGATGCTTCTTCTTAACTAACCAATACAGGCGCACCGCTGG 1980
 Qy 1986 GGTCAATGTTGCGGCTGACGCTGATGAACTCTTCTGGGTACACCAAGACTTGGCGGCGAC 2045
 Db 1981 GCAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2040
 Qy 2046 CACCTGCGGCTATGAGAGTGTCTTCAACCGGACGACGAGCTTGTGTTGTTGTTGTTG 2105
 Db 2041 CTTCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2094
 Qy 2106 GTTTAGGAGGATCTGATACCTTACCTCAATGCGGCTCTGGGCGCTGCTGCTGCTGCTG 2165
 Db 2095 GCTTCCGCAAGATCCGACGCTCATCTCTCGGTGCGGCTCCGCTCCCTGATGATGATG 2154
 Qy 2166 CAAGGTGCTGATGCACTTACCTTACGCTTCTGCGATTTACCTTCTGACAGTTTAACTATA 2225
 Db 2155 CAGAGTGGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2214
 Qy 2226 CCATCTTCAAAATAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2285
 Db 2215 CCATATTTAAATCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2274
 Qy 2286 ATTTTCACTGTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2345
 Db 2275 ACTGGACGCGGGCGAAGCTTGGGATGATGATGATGATGATGATGATGATGATGATGATG 2334

QY	2346	TGTTGACATCCACACGAATGGCCATTTTACCTTCTTACTCGGACCTCCCGCCT	2405	Db	3415	TGTCGCGCCCATCATCGCGGTACGCCAGACAAAGGCGCTCTCTAGGGTGCAATCA	3474
Db	2335	TACTGTGACCATACACAGTGGAGGTCTCCGCTGTTCTTCAACAACCTTACAGCCT	2394	QY	3486	TGAGCATGACGGGGCGGACAGACAGACAGAGCTGGGAAATTCAGGTCTCTGCACAG	3545
QY	2406	TGTGCACTGTCTTCTCCACCTCCACAAACATCTGGGACGTAACAATTCATATAGGCC	2465	Db	3475	CCAGCTTAATCTGCGCGGACAAACAAAGTGGAGGTGAGGTCTGATTTGTCAACTG	3534
Db	2395	TGTCACCGGCTCATCCACCTCCACAGAACTTGTGGACGTGCACTTGTACGGGG	2454	QY	3546	TCACCTAGTCTCTCGGAAACATCATCTCGGGGTCTTGTGGACTGTCTACCATGGAG	3605
QY	2466	TATCACCCTGCCCTACAAATATACATCGTCGATGGGAGTGGTAATACTTATCTCTGC	2525	Db	3535	CTGCCAAACCTTCTTGGCAACGTGATCAATGGGGTGTCTGACTGTCTACACGGGG	3594
Db	2455	TGGGGTCAAGCATCGCGCTCTCGGCCATTAAGTGGGAGTACGCTCTCTCTCTC	2514	QY	3606	CTGGCAACAGACTCTGGCCGCTCACGGGGTCCGGTCAACGAGATGTACTCAAGTGTG	3665
QY	2526	TCTTAGCGGACGCGAGGTCTGCGCTCTTATAGGATGCTCATCTTGTGGGCGAGCGCG	2585	Db	3595	CCGGAACAGAGGACCATCGCTCACCAAGGCTCTGTCAATCCAGATGTATACCAATGTAG	3654
Db	2515	TGCTTGACAGCGCGGCTCTCTCTGCTTGTGGATGCTACTCATATCCAGCGG	2574	QY	3666	AGGGGACTTAGTGGGTGGCCAGCCCTCGGACTTAAATCTTGGAGCGTGCACCT	3725
QY	2586	AAGCAGCATAGAGAGCTGGTCACTCTGACGCTCGAGCGAGCTAGTGTCAATGGCT	2645	Db	3655	ACCAAGACCTTGTGGCTGGCCGCTCGCAAGAGTAGCGCTCATTTGACACCTGACCTT	3714
Db	2575	AGCGGCTTTGGAGAACCTCGTAATCTTAATGAGCATCCCTGGCGGAGCACGCGTC	2634	QY	3726	GTGGAGCGGTGACCTGTACTGTGTCACGCGAAAGCTGATGTCACTCCCGGCTCGAAGAC	3785
QY	2646	TCTTATATTTGTCATCTTTTCTGCTGCTTGTGTGATCATCAAGGTCGGTGTAGTCCCT	2705	Db	3715	GCGGCTCTCGGACCTTACTTACTGTGTCAAGGACGCGCATGTCTTCCGTGGCGCGC	3774
Db	2635	TGTATCTCTCTCGTGTCTCTCTGCTTGTGATGGTATTTGAAGGGTAAGTGGGTGCCG	2694	QY	3786	GCGGGGACAAACGCGGAGCGCTACTCTCCCGAGACCTCTTTCACCTTTGAAGGGTCTCT	3845
QY	2706	TAGTACTATCTCCCTCACTGGGCTGTGTCTCTTTAGCCTTACTGCTCTAGCATTTGCCCT	2765	Db	3775	GGGTGATACAGAGCGGACAGCTGTCTCGCCCGGCGCATTTCTACTTTGAAAGGCTCTCT	3834
Db	2695	GAGGGTCTACACCTTCTACGGATGTGGCTCTCTCTCTGCTCTCTTGGCGTGGCCCT	2754	QY	3846	CAGGAGCGGCTGTATGCCCGGAGCGGACGCTGTGCGAGTCTTCCGGGAGCTGTCT	3905
QY	2766	AACAGGCTTATGCTATGACGATCTGTGATGCGGACGATAGAGCGGCTCTCTCGGTAA	2825	Db	3835	CGGGGGTCTCGGTGTGTGTCGCCCGGGGACGCGGTGGGCATATTTAGGGCGCGGTGT	3894
Db	2755	AGCGGCGTACGCGCTGACACGAGGTGCGCGCTGCTGTGCGGTGTGTCTCTGTCT	2814	QY	3906	GCTCTCGGGCTGGCTAAGTCCATAGATTTTCAATCCCGCTTGAGACACTCGACATCGTCA	3965
QY	2826	TGATCACTCTTACTCTCACCCCGGCTATAGACCTTCTCAGCGGCTTTTGTGGT	2885	Db	3895	GAACCGTGGAGTGGCTAAGGCGGTGGACTTTTATCCCTGTGGAGAACCTTAGACAAACA	3954
Db	2815	GTTTGTGGGCTGACTCTGTCAACATTTTACAAGCGCTATATCAGTGGTGTGTGGT	2874	QY	3966	CAGCGTCCCCACCTTTTGTGACAAACAGCACCACTCTCTGTGCCCCAGACCTATCAGG	4025
QY	2886	GTTTGTGCTATCTTCAACCTCGGGGAGCTATGCTCAGAGTGGGACCACTATGTC	2945	Db	3955	TGAGGTCTCCCGGTGTTCAGGATACTCTCTCCACAGTAGTGTGCCAGAGCTTCCAGG	4014
Db	2875	GGCTTCAATTTTCTGACAGAGTGAAGGCACTGACAGTGTGGATTCCCGCCCTCA	2934	QY	4026	TGGGTACTTGTGATGCCCGGACTGGCAGTGGAAAGAGCACCAAGTCTCTGTCCGATATG	4085
QY	2946	AGTGTGCGGTGGCGGTGATGGCATCATATGGGCGGTGCGCATATTTCTACCCAGGTGTG	3005	Db	4015	TGGCTCACCTCTCATGCTCCACAGCAGCGGCAAAAGACCAAGGTCTCCGGTGTGATATG	4074
Db	2935	ACGTCCAGGGGGCGGCGCGCTCATCTTACTGTGTGCTGTACACCGGACTCTGG	2994	QY	4086	CTGCTCAGGGGTATAAGTGTAGTGTCTTAACTCCCTCAGTGTGCTGCACTCTGGGGTGTG	4145
QY	3006	TGTTTGTACATAACAGTGTCTCTTGGCGGTCTTGGCGCTGTACTCTCTTAAAGGTG	3065	Db	4075	CAGCTCAGGGCTATAAGTGTGTACTACTCAAACCTCTGTGTCTGCAACACTGGGGTGTG	4134
Db	2995	TATTTGACATCAACAAATGCTGTGCGGCTCTTGGACCCCTTGGATTCTTCAAGCCA	3054	QY	4146	GGCGTACTTGTCTAAGGCATGTGCATCAATCCCAACATTAAGGACTGGAGTCAGGACTG	4205
QY	3066	CTTTGACGCGCTGCTGCTTCTGTCAGGGCTCACGCTCTACTGAGGATGTGACCATGG	3125	Db	4135	GTGCTTACATGTCTAAGGCTCATGGATCGATCCTTAACATCAGGACCGGGGTGAGAACAA	4194
Db	3055	GTTTGTCTTAAGTACCTTCTTGTGCGCTTCCAGGCTTCTCCGGTCTGCGCGTTAG	3114	QY	4206	TGACGACCGGGGCGCCATCAGTACTCCACATATGGCAATTCCTCGCGATGGGGCT	4265
QY	3126	CAAGGCATCTGCGGGGGGAGGTACGTCAGATGTGGCTTACTAGCCCTTGGCAGGTGGA	3185	Db	4195	TTACCACTGGCAGCCCATCAGTACTCCACCTACGCGAAGTTCCTTTCGCGAGCGGGT	4254
Db	3115	CGCGAAGATGATCGAGGCGCATTTACGTGCAAAATGGTTCATTAAGTTAGGGCGCGTTA	3174	QY	4266	GTGCGGGCGGCTACGACATCATCATATGTGATGAATGCAATGCGCTGCGGTCTTACCA	4325
QY	3186	CTGGCACTATCATATGACCACTCACCCCTATGTGCGATGGGCTGTAGTGGCTGTC	3245	Db	4255	GCTCGGGGGCGCTTATGACATAATATTTGTGACAGTGGCCACTCCACGGATGCCACAT	4314
Db	3175	CTGGCACTATGTTTAAACATCTCACTCTCTTGGGACTGGGCGACAAACGGCTGTC	3234	QY	4326	CCATCTTGGCATCGGAAACAGTCTTGTATCAAGCAGAGACAGTGGGGTCAAGTAACTG	4385
QY	3246	GGGACCTGGCGGTGCGCTTGTAGGCTATCATCTTTCAGTCCGATGGGAGAAAGTCAATTG	3305	Db	4315	CCATCTTGGGCATCGGCACTGTCTTGACAAAGCAGAGACTGCGGGGGCGAGACTGGTGTG	4374
Db	3235	GAGATCTGGCGGTGTAGAGCCAGTCTCTTCTCCAAATTTGGAGACCAAGCTCATCA	3294	QY	4386	TGCTGTACTAGCTAGCCCCCTGGGTCAGTGACACCCCGCCACCATACATAGAGGAGG	4445
QY	3306	TCTGGGAGCGGAGACAGTCTTGTGGGACATTTTACAGGACTTCCCGTGTGCGCC	3365	Db	4375	TGCTGCCACCGCCACCCCTCGGGGTCTCGTCACTGTGCCCCCATCCCAACATCGAGAGG	4434
Db	3295	CGTGGGGGAGATACCGCGCGTGTGCTGACATCATCAACCGCTTCTCTTTCGCGCC	3354	QY	4446	TGSCCTTGGCAGGAGGCGGAGATCCCTTCTATGGAGGGCGATTTCCCTCTCTTACA	4505
QY	3366	GACTTGTCTGGAGGTCTCTTGTGGCCAGCTGATGCTATACCTCCAGGGGTGGATC	3425	Db	4435	TTGCTCTGTCCACCGGAGAGATCCCTTTTACGGCAGGCTATCCCCCTCGAAGTAA	4494
Db	3355	GCAGGGCCCGGAGATCTGCTGGGCGCCGCGATGGTCTTCCAGGGGTGGAGGT	3414	QY	4506	TCAAGGGAGGAGACATCTGATCTTCTGTGCAATTTCAAAGAAAAAGTGTGACGACTCGCGG	4565
QY	3426	TTCTCGCCCCCATCACTGCTTACGCCCGCAGACACGCTGGCGCTTTTGGGACCATAGTGG	3485				

Db	4495	TCAAGGGGGGAGACATCTCATCTTCTGTCAATCAAGAAAGATGCGAGAACTCGCCG	4554	QY	5646	TGTGGAACCTTCATTTAGCGGCATCCAAATACCTCGCAGCACTATTCAACACCTGCCAGGGAACC	5705
QY	4566	CGGCCCTTCGGGGTATGGCTTGAACCTCAGTGGCATACTACAGAGGGTGGAGTCTCCG	4625	Db	5635	TGTGGAACCTTCATTCAGTGGGATACAAATCTTGGCGGGCTTGTCAACGCTGCCTGGTAACC	5694
Db	4555	CAAAAGCTGTGCGAATGGGCAATCAATCGCGTGGCTTACTACCGGGTCTTGAAGTGTCCG	4614	QY	5706	CTGCACTAGCTTCCATGATGGGGTTCAGTGGCGCCCTCACCACTCGCGCTGTCAACAAGCA	5765
QY	4626	TAATACCAACTCAGGGAGAGCTAGTGGTCTGCGCACCGAGCGCCCTCATGACAGGGTATA	4685	Db	5695	CGGCCATTCCTTCATTTAGTGGCTTTTACAGCTGCTGTCAACAGCCACCTAACCACTAGCC	5754
Db	4615	TCATCCCGACCGAGCGGCGATGTTGTCTGTGGCAACCGATGCCCTCATGACCGCTATA	4674	QY	5766	CACTATCTCTTCAACATTTTGGGGGGCTGGGTAGCTAGCATCCCAATTTGCACCAACCGCGG	5825
QY	4686	CTGGGAGCTTGGCTCGTATCGACTGCAAGCTGCAAGCTGCTCAAGCTGCTCAAGCTA	4745	Db	5755	AAACCCCTCTCTTCAACATTTTGGGGGGGTGGGTGGCTGCCAGCTCGCGCCCGCGGTG	5814
Db	4675	CCGGGAGCTTCGACTCGGTGATAGACTGCAATACGTGTGTGTGTGTGTGTGTGTGTGTGT	4734	QY	5826	GGCGCACTGGCTTGT	5885
QY	4746	GTTTAGACCCCAATTCACCAATACCAACAGATTTGCTCCCTCAAGAGCTGCTCAAGCTA	4805	Db	5815	CCGCTACTGCTTGTGTGGGGCTGGCTTAGCTGTGGCGCCGCGCATCGGAGTGTGGAGTGG	5874
Db	4735	GCCTTGACCTTACCTTCAACATTTGAGACAATCAGCTCCCGCAGGATGCTGTCTCCGCA	4794	QY	5886	GTAAGTGTGTGTGGGATCTCTGGGAGGGTATGGTGGGGGCAATTTCCGGGGGCTCTCGTGG	5945
QY	4806	GCCAGCGCCGGGTGCGACGGGTAGGGAGAGCTGGGCAATTTATAGGTATGTTTCCACTG	4865	Db	5875	GGAGGTCTCTATAGACATCTTGCAGGGTATGGCGCGCGCGTGGCGGAGCTCTTGTGTG	5934
Db	4795	CTCAACGTGGGGGAGGACTGGCAGGGGGAAGCAGGCATCTACAGATTTGTGGCACCGG	4854	QY	5946	CATTCAAGATCATGTCTGGGAGAGCCCTCCATGAGAGATGTGCTCAACTTCTGCTGCTG	6005
QY	4866	GTGAGGAGCTCAGGAATGTTTGAAGTGTAGTGTCTGTGTGTGTGTGTGTGTGTGTGTGT	4925	Db	5935	CATTCAAGATCATGAGCGGTGAGTCCCTTCCACGGAGGACCTGGTCAATCTACTGCTCC	5994
Db	4855	GGAGGCGCCCTCCGGCATGTTGCACTCGTCTCGTCTGTGTGTGTGTGTGTGTGTGTGTGT	4914	QY	6006	GAATCTGTCTCGGGTGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	6065
QY	4926	CCGATGTTATGAGCTCACACCATCGAGACACCGTCAAGCTCAGGGCTATTTCAACA	4985	Db	5995	CCATCTCTGCGCGGAGCCCTGCTAGTGGCGGTGTCTGTGAGCAATACTGCGCGCGG	6054
Db	4915	GTGCTTGTGTATGAGCTCACCGCGCGAGACTACAGTTAGGCTACGAGCGTACATGAACA	4974	QY	6066	ACGTGGGACCGGGGAGGCGCGCTCAATGGATGAATAGACTCATTTGCTTTGCTTTCCA	6125
QY	4986	CGCCCGGTGTGCTGTGCAAGACCATCTTGTGATTTTGGGAGGAGTGTTCACCGGCC	5045	Db	6055	ACGTTGGCCCGGCGAGGGGCGAGTGCAGTGGATGAACCGGCTGATAGCTTCGCTCCC	6114
Db	4975	CCCCGGGGTTCCTGTGTGCGAGACCATCTTGAATTTTGGGAGGGGCTTTTACAGGCC	5034	QY	6126	GAGGAATCAGCTGCGCCCCCACCCTACGTACGTACGAGGTGCGGTGCGACGCTGTGA	6185
QY	5046	TCACACATAGATGCCCTCTCTTTCCTTCCCAACAAAGCAATCGGGGAAAAATTTCCGAT	5105	Db	6115	GGGGGAACCATGTTTCCCGCCACCGCATACGTGCGGAGCGCATGCACTGCGCGGTCA	6174
Db	5035	TCATCATATAGATGCCCTCTCTTATCCAGACAAAGACAGAGTGGGGGAACTTCTCTT	5094	QY	6186	CCCACTACTTGGCTCCCTTACCATAAACAGCTGCTCAGAGACTCCCAACCTGATTA	6245
QY	5106	ACTTAAACAGCTACAGGCTACAGTGTGCGCTAGGGCCAAAGCCCCCCCCCTGCTGGG	5165	Db	6175	CTGCCATCTCAGAGCTCTACTGTAAACAGCTCTGAGGCGACTGCACTAGTGTGATTA	6234
Db	5095	ACCTGTAGGCTACCAAGCAGCGTGTGCGCTAGGGCTCAAGGCCCTTCCCGCATCTGGG	5154	QY	6246	CTGAGGACTGCCCATCTCCATGCGGGGCTGCTGGCTCCCGCATGTGTGGGACTGGGTTT	6305
QY	5166	ACGTATGTGGAAAGTGTGACTCGACTCAAGGCCCACTCGTGGGCCCCCACCCTCTCC	5225	Db	6235	GCTCGAGGTGTAACCACTCCATCTCGGTTCCTGGCTAAGGGACATCTCGGACTGATAT	6294
Db	5155	ACCAGATGTGGAAAGTGTGATTTGCTCGCTCAAGCCCCACCTCCATGGGCCAACCCCTGC	5214	QY	6306	GCACCATCTATACAGACTTTTAAATTTGCTGACCTCCAAATTTATTTCCCAAGATGCCG	6365
QY	5226	TGTACCGCTTGGGCTGTGTACAAACAGAGTCACTTCAACATCTCGGTGACGAATATA	5285	Db	6295	GGAGGTGTGAGCGGACTTTAAGACCTTGGCTAAAGCTAAGCTCATGCGCACAGCTGCCCTG	6354
Db	5215	TATACAGACTGGGCGTGTTCAGAAATGAAATCAACCTGACGCAACCCAGTCAACCAATA	5274	QY	6366	GCTCCCTCTTGTCTCTCTCAAAAGGGGTACAAGGGCGTGTGGGCGCGCATCTGGCATCA	6425
QY	5286	TCGCCACCTGCATGCAAGCGGACTTGGAGTCAATGACAGCAATGGGCTTTCGCGAGGGG	5345	Db	6355	GGATCCCTTGTGTCTCTCGAGCGGGGTATAGGGGGTCTGGCGAGTGGACGGCATCA	6414
Db	5275	TCATGACATGATGTGCGCGGAGCTGGAGTCTGACGAGCACTGGGTCTGTGGGG	5334	QY	6426	TGACCAACAGCTGTCTTGTGCGGCGCAATATCTCTGGCAATGTCTGCGCTTGGGCTCATGA	6485
QY	5346	GAGTCTTGGCGGCGCTCGCCGCTATTGCTGGCGAGCGGGTGTGTTGTCATCATCGGCC	5405	Db	6415	TGCACACTCGCTGCCACTGTGGAGCTGAGATCACTGGACATGTCAAAAGGCGGACGATGA	6474
Db	5335	GGGTCTGGCTGTGTTGGCGCGTATTGCTGTCAACAGGCTCGGTGGTCAATAGTGGCA	5394	QY	6486	GAAATCAGCGGGCTTAAGACTTCATGAAATATCTGGCAGGGGACCTTTCTCTATCAATGTT	6545
QY	5406	GCTTGCATTAACACAGGAGCGCTGTGCTGGCGGCAAGAGAGTCTCTATGAGGCTT	5465	Db	6475	GGATCGTCTGCTTAGGACCTTGAAGCAATGTGGAGTGGGACCTTTCCCATTAATGCTCT	6534
Db	5395	GGGTCTGTGTTGTCGGGAAGCGGCAATCATCTGACAGGGAAGTCTCTACCGGAGAT	5454	QY	6546	ACACGGAGGGGCGAGTGGCTGCGGCAACCGCGCCCAAACTTTAAGGTGCGCATCTCGAGGG	6605
QY	5466	TTGATGAGATGAGGATGTGCTCTAGGGCGCTCTATTGAAGAGGGGACCGGATAG	5525	Db	6535	ACACACGGGGCGCTGTATCCCGCTTCTGCGCGGAACTACAGTTTGGCGCTATGGGGG	6594
Db	5455	TCGATGAGATGGAAGAGTGTCTCAGCATTACCGTACATCGAGCAAGGGGATGATCTCG	5514	QY	6606	TGGCGGCTCTCAGAGTACGCGAGGTGACGCGACACGGGTCAATCACTATATAACAGGAC	6665
QY	5526	CCGAGATGCTGAAAGTCCAAGATCCAAGGCTTATTGACAGCAAGCTTCCAAACAGCTCAAG	5585	Db	6595	TGTCTCAGAGGAAATATGTGGAGATAAGGACGAGTGGGGGACTTCCACTAGGTGACGGGTA	6654
Db	5515	CCGAGCAGTTCAAGCAGAGGCGCTCGGCTCTCTGACAGCGGCTCCGCTCAGGAGAGG	5574	QY	6666	TCACCATGATTAATTTGAAGTCCCTTGCCTAACTACCTCTCTCCGAGTCTTTTCTCTGGG	6725
QY	5586	ACATACAAACCTGTGAGGCTTTCATGCGCCCAAGGTAGAACAAATTTCTGGGCAACACA	5645	Db	6655	TGACTACTGACAACTCTCAAAATGCGCGTGCAGGTCCCATCGCCCGAAATTTTTCACAGAAT	6714
Db	5575	TTATCGCCCTGTGTCCAGACCAACTGGCAAAACTCGAGACCTTCTGGGCGAAGCAT	5634				

QY 6726 TGACGAGGTGACATCCATAGTTTGGCCCCCAGACCGAGCGTTTTCGGGATGAGG 6785
 Db 6715 TGACGGGTGCGCTACATAGTTTGGCCCCCAGACCGCTTGGCTGGGGAGAGG 6774
 QY 6786 TCTCGTTCTCGGTTGGCTTAATTCATTTGCTGCTGGGTCCGAGCTTCTTGGAGCCCTG 6845
 Db 6775 TATCATTTAGAGTAGACTCCAGAAATACCGGTAGGGTGCATATACCTTTGGAGCCCG 6834
 QY 6846 AACCCGACACAGAGTATTCATGCTCCATGCTTAACAGATCCATCTCATATACCGCGCGAGA 6905
 Db 6835 AACCGAGCTGGCGGTGACCTTCCATGCTTACATCCCTCCCATATACAGCAGAGG 6894
 QY 6906 CTGACGCGCGCTTTAGCGGGGGTCAACCCCATCCGAGCAAGCTCTCCGCGAGCC 6965
 Db 6895 CGCGCGCGCAAGTTGGCGAGGGATCAACCCCTCTGTGGCAGCTCTCCGCTAGCC 6954
 QY 6966 AGCTATCGGCACCATCGCTCGAGCCACCTGACACCAACCGCAAGCTTATGATGTGG 7025
 Db 6955 AGCTATCGGCTCCATCTCTCAAGGCAATTTGACCGCTAACCATGATCTCCCTGATGCTG 7014
 QY 7026 ACATGGTGATCTAACT-----GTTTCATGGGGGGGATGATGACTCGATAG 7073
 Db 7015 AGCTCATAGAGGCGCACTCTATGAGGAGGAGATGGGCGCAACATCACAGGGTTG 7074
 QY 7074 AGTCTGGGTCCAAAGTGGTCTGACTCTCTCGACCCCAATGGTCCGAAAGAGAGCG 7133
 Db 7075 AGTCAGAAACAAGTGTGATTTCTGACTCTCTTGNATCGCTTGTGGCGGAGGAGCG 7134
 QY 7134 ACCTTGAGCTTCGATACCATCAGAAATPACATGCTCCCAAGAGAGGTTCCCAACAGCTT 7193
 Db 7135 AGCGGAGATCTCCGTACCGCGCAAAATCTCGGGAATCTCGGAGATTCGCGCAGGCC 7194
 QY 7194 TAGCGCTGGGCAACGCTGATTAACCAACCGCTTGGATCTGGTGAAGAGAGCGGAG 7253
 Db 7195 TGCCCGTTGGGCGCGCGGATATTAACCCCGCTAGTGGAGAGCTGGGAAAGAGCCG 7254
 QY 7254 ATTACCAACCGGCACTTGTGGGCTGTGCTCTCCCTCTCTAGGAAACCCGACGC 7313
 Db 7255 ACTAGCAACACCTGTGCTCATGCTGCTCGCTTCCACTCCAAAGTCCCTCTGTCG 7314
 QY 7314 CTCCTCCAGAGGCGCGGACAGTGGGCTTAAGTAGAGACTCCATAGAGAGTGCCTTC 7373
 Db 7315 CTCGCTCGGAAGAGCGAGCTGTGCTCTCAATCAACCTATCTACTGCTGCTGG 7374
 QY 7374 AACAGCTGGCCATTAAGTCTTTGGCAGCGCCCCCCCCAAGCGGATTCAGGCTTTCCA 7433
 Db 7375 CCGAGCTCGCCACAGAAAGCTTTTGGCAGCTCTCACTTCGCGCATTTACGGGCGCAATA 7434
 QY 7434 CGGGGCGGCGCTGCGGATTCGCGCAGTCCAGACGCTCTCTGATGAGTTGGCCCTTTCCG 7493
 Db 7435 CGACCAATCTCTG-----AGCCCGCCCTTCTGGCTGCCCCCGACTCCG 7482
 QY 7494 AGACAGGTTCCATCTCTTCCATGCCCCCTCGAGGGGAGCTTGGAGATCCAGACTGG 7553
 Db 7483 ACGCTGAGTCTATTCTCCATGCCCCCTCGAGGGGAGCTTGGGATCCGATCTTA 7542
 QY 7554 AGCCTGAGCAGGTAGAGCCCCCCCCCCCCCGGAGGCTGCGAGCTCCCGGCTCGG 7613
 Db 7543 -----GGAGC 7548
 QY 7614 ACTCGGGTCTGCTACTTGTCTCGAGGAGGAGACTCCGCTGCTGCTGCTCTCATGT 7673
 Db 7549 GGTATGCTCAACGGTCACTAGTGAGGCGCAACGGGAGGATGCTGCTGCTCTCATGT 7608
 QY 7674 CATACTCTGAGACCGGGCTCTAATACTCTTGTAGTCCGAGAGGAGGATTCACCGA 7733
 Db 7609 CTTACTCTTGACAGGCGCACTGCTCACCCGCTGCGCGGGAAGAACAGAAATGCGCCA 7668
 QY 7734 TTAACCCCTTACCACTCTCTGTTGGATATCACACAGAGTGTACTGTACCAACAA 7793
 Db 7669 TCAATGACACTAAGCAACTGCTGTACTAGTCAACCAATTTGGTGTATTCACCACTCAC 7728
 QY 7794 AGAGCGCTCACTAAGGGCTAAAAAGGTAACTTTTGTAGATGCAAGTGTCTGACTCTCT 7853

Db 7729 GCAGTGTCTTGCACAAAGGAGAAAGTACATTTTGACAGACTGCAAGTTCTTGGACGCG 7788
 QY 7854 ACTTAGACTCACTCTTAAGGACATTAAGCTAGCGCTCCCAAGGTTCACGCAAGCTCC 7913
 Db 7789 ATTACGAGACGTACTCAAGGAGGTTAAAGCAGCGGCTCAAAAGTGAAGGCTTAATTGC 7848
 QY 7914 TCACCATGAGGAGGCTTGCAGTTTAAACCCCAACCCCATTTCTGCAAGATCTAATATGGGT 7973
 Db 7849 TATCGTAGAGGAGCTTGCAGCTTCAGCCCTCAGCCCTCAGCTCAGCAAAATCCAAGTTTGGTT 7908
 QY 7974 TTGGGGCTAAGGAGTCCGACGCTTGTCCGAGGCGCTTAAACCAATCAAGTCCGCTGT 8033
 Db 7909 ATGGGCAAAAGACGCTTCTGCTATGCGGAGGCGCTAAACCCCAATCAATCCGCTGT 7968
 QY 8034 GGAAGGACCTCTGAGGAGCTCAGAAACACCAATTTCCCAACCAATTTATGGCCAAAAATG 8093
 Db 7969 GGAAGGACCTTCTGGAAGCAATTAACCAACCAATAGACTACCTATCATGGCTAAGACG 8028
 QY 8094 AGTGTCTGCTGAGACCCACCAAGGGGGGCAAGAAAGCAGCTCGCTTATCGTTTACC 8153
 Db 8029 AGTGTCTGCTGAGCTTCAAGCTTGAAGGGGGCTGTAAGCCAGCTCTCATCGTTTCC 8088
 QY 8154 CTGACCTCGGCTGAGGCTCTCGAGAGATGGCCCTTTTATGACATTAACAAAACTTC 8213
 Db 8089 CCGATCTGGGCTGCGCTGTGCAAAAGATGGCTTGTACGACGTGGTTTACAAAGCTCC 8148
 QY 8214 CTCAGCGGTGATGGGGCTTCTTATGATTTCCAGTATTTCCCGCTCAGCGGTAGAGT 8273
 Db 8149 CTTTGGCGGTGATGGGAAGCTCTTACGATTTCAATACTCAGCAGGACAGCGGTTGAAT 8208
 QY 8274 TTCTCTTGAAGCATGGGCGGAAAGAGACCTATGCTGTTTTCGTATGATACCCGAT 8333
 Db 8209 TCTCTGCAAGGCTGGAAGTCCAGAAACCCCAATGGGTTCTGATATGATACCCGCT 8268
 QY 8334 GCTTTGACTCAACCGCTCACTGAGAGAGACATCAGGACTGAGGAGTCCATATATCGGGCT 8393
 Db 8269 GCTTTGACTCCACAGTCACTGAGAGGAGATCCGTACGAGGAGGCAATCTACCAATGTT 8328
 QY 8394 GCTCTTGGCGGAGGAGGCGCACCTGCCATACCTCGCTTAAGTACAGAGACTTACGTTG 8453
 Db 8329 GTGACCTCGACCCCAAGCCCGGTGCCATCAATGCTCTCAGCAGAGAGCTTATGTTG 8388
 QY 8454 GAGGGCTTATGTTCAACAGCAAGGGCAACCTTCGCGGTACAGGCTTTCGCGCGCAGCG 8513
 Db 8389 GGGGCTCTTACCAATTCAGGGGGAGAACTGCGGCTATCGAGGTGCGCGCAGCG 8448
 QY 8514 GGTGCTCACTAGCTAGGAGGAGAACCATCATCATGCTACGTGAAGGCTTACGCGCTT 8573
 Db 8449 GGTACTGACAACTAGCTGTGGTAAACCTCACTTGTCTACATCAAGGCGCGGCGAGCT 8508
 QY 8574 GTAAAGCTGAGGGATAATCGCGCCCAATGCTGTTGATGCGCGGATGACTTGGTTGTCA 8633
 Db 8509 GTGAGCGCGAGGCTTCCAGGACTGCAACATGCTGCTGTGTGGCGACGACTTATGCTTA 8568
 QY 8634 TCTCAGAAAGCAGGGGAGCCGAGGAGAGAGCGGAAACCTGAGAGCTTTCACGAGGCTA 8693
 Db 8569 TCTGTGAAGCGCGGGGCTCCAGAGGAGCGCGGAGCTTCCAGAGCTTTCACGAGGCTA 8628
 QY 8694 TGACAGGTTATGTCCTTCTGTTGAGCCCCCCCCAGACGAGGATGATCTCGAGCTGA 8753
 Db 8629 TGACAGGTTACTCGCGCCCCCTGGGAGCCCCCAACAGAAATACGACTTGGAGCTCA 8688
 QY 8754 TAACTCTTCTCTCAATGTCTGTGGCGTGGGCGCCCAAGCGCGCGCAGATACT 8813
 Db 8689 TAACTATCTCTCTCAAGTGTCTAGTGTGCTGCGGAGGCTGCGTGGGAGACAGCAAGAC 8748
 QY 8814 ACCTGACAGAGACCTTACACTTCAATCGCCCGGCTGCTCGGAAACAGTGTAGACAT 8873
 Db 8749 ACCTTACCGGTGACCTTCAACCCCCCTCGGAGAGCTGCGTGGGAGACAGCAAGAC 8808
 QY 8874 CCCCTGTCAATTCATGGCTGGGAAACATCATTCAGTACGCGCCCGACATATGGGCTCGCA 8933

8809 CTCAGTCAATCTCTGGCTAGGCAACATAATCATGTTTGGCCCCACACTGTGGGCGGAGG 8868
8934 TGGTCTGTGATGACACACTTCTTCCATTCTCATGGCTCAAGACAGCTCGACGACCAACC 8993
8869 TGATCTGATGACCACTTCTTGTAGCTCTTATAGCCAGGACCGCTTGAACAGGCC 8928
8994 TCAACTTTGAGATGTACGAGCGGTGTACTCCGTGAGTCCCTTGGACCTCCAGCTATAA 9053
8929 TCGATTGCGAGATCTACGGGCGCTTACTCCATAGAACCACTTGATCTACTCCCAATCA 8988
9054 TTGAAGTTTACATGGCTTGAACGCTTTTCTCTGACACATACACTCCCAAGAACTGA 9113
8989 TTCAAGACTCCATGGCTCAGCGCAATTTTCACTCCACAGTTACTCTCCAGGTGAATTA 9048
9114 CACGGTGGCTTCAGCCCTCAGAAACTTGGGGCGCCACCCCTCAGAGCGTGGAGAGCC 9173
9049 ATAGGTTGGCGCATGCTCAGAAACTTGGGTACCGCTTGGAGCTTGGAGACACC 9108
9174 GGGCAGCTGAGTCCAGGCGCTCCCTCATCTCCGTGGGGGAGAGCGCGCTTGGCGTC 9233
9109 GGGCGCGGAGCTCCGCGTAGGCTTCTGGCAGAGGAGCGGCTGCCATATGTGCA 9168
9234 GATATCTTCAATTTGGGGGTGAAGACCAAGCTCAAACTCACTCCANTGCCGGAAGCGC 9293
9169 AGTACCTCTTCAACTGGGCGAGTGAAGAACAAAGCTCAAACTCACTCCAAATAGCGCGCTG 9228
9294 GCCTCTGGATTATCCAGCTGTTTACCGTCCGCGCGCGGGGCGACATTTATCACA 9353
9229 GCCAGCTGGACTTGTCCGCTGTTTACCGTGTGCTACGGTGGGAGACATTTATCACA 9288
9354 CGGTGTCGCTGCCCGACCCCGCTTATGCTCTTTGGCTTACTCTACTTTTGTAGGG 9413
9289 GCGGTGCTCATGCCCGCGCGCTGATCTGTTTGGCTTACTCTCTGTTGTCAGGG 9348
9414 TAGCCCTTTTCTACTCTCCGCTCGGTAGAG 9444
9349 TAGGCATCTACTCTCTCCCAACCGATGAAG 9379

RESULT 15
US-08-440-549-9
; Sequence 9, Application US/08440549
; Patent No. 6312889
GENERAL INFORMATION:
APPLICANT: Houghton, Michael
APPLICANT: Choo, Qui-Lim
APPLICANT: Kuo, George
TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: P.O. Box 8097 (Int. Prop. R-440)
CITY: Emeryville
STATE: CA
COUNTRY: U.S.A.
ZIP: 94662-8097
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,549
FILING DATE: 12-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/910,760
FILING DATE: 07-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Blackburn Esq., Robert P.
REGISTRATION NUMBER: 30,447
REFERENCE/DOCKET NUMBER: 0101.002

TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2702
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 9401 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 342..9374
FEATURE:
NAME/KEY: misc_feature
LOCATION: 366
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Arg."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 372
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 867
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1341
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2148
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ile."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2883
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Asn."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3681
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3690
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4167
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Leu."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4323
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4701
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Tyr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4752
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."

```

FEATURE:
NAME/KEY: misc_feature
LOCATION: 5970
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6193
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be His."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6186
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Cys."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6402
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7386
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7494
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Phe."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7497
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ala."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7845
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Phe."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 8409
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9102
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also by Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9327
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Pro."

```

US-08-440-549-9

```

Query Match 45.6%; Score 4425.8; DB 4; Length 9401;
Best Local Similarity 67.3%; Pred. No. 0;
Matches 6364; Conservative 0; Mismatches 2997; Indels 90; Gaps 5;

Qy 6 CCCCTAATAGGGGACACTCCGCGATGATCACTCCCTGAGGAACTACTGTCTTCA 65
Db 7 CCCCTGATGGGGGACACTCCACATGATCACTCCCTGAGGAACTACTGTCTTCA 66
Qy 66 CGCAGAAAGCGTCTAGCCATGGCGTGTAGTATGATGTCGTACAGCCTCCAGGCCCGCCCC 125
Db 67 CGCAGAAAGCGTCTAGCCATGGCGTGTAGTATGATGTCGTACAGCCTCCAGGCCCGCCCC 126
Qy 126 TCCCGGAGAGCCATAGTGTGTCCGGAACCGGTGAGTACACCGGAATTCGCGGAGAGAC 185
Db 127 TCCCGGAGAGCCATAGTGTGTCCGGAACCGGTGAGTACACCGGAATTCGCGGAGAGAC 186

```

```

186 TGGGTCTCTTTCTTGGATAAACCCACTCTATGCCCCCGCCATTTGGGGTGGCCCCGCAAGA 245
187 CGGGTCTCTTTCTTGGATCAACCCGCTCAATGCTGGAGATTTGGGGTGGCCCCGCAAGA 246
246 CTGCTAGCCGAGTAGCGTTGGTTGCGAAAGCCCTTGTGGTACTGCTGATAGGGTCTTT 305
247 CTGCTAGCCGAGTAGTGTGGGTGCGAAAGCCCTTGTGGTACTGCTGATAGGGTCTTT 306
306 GCGAGTGGCCCCGGGAGGTCTCTAGACCGTGCACCATGAGCACAATCTTAACCTCAA 365
307 GCGAGTGGCCCCGGGAGGTCTCTAGACCGTGCACCATGAGCACAATCTTAACCTCAA 366
366 GAAAAACCAAAAAACACCAACCGTGCACCACAAGAGCGTTAAGTTTCCGGCGCGCGCC 425
367 AAAAAACCAAAAAACACCAACCGTGCACCACAAGAGCGTTAAGTTTCCGGCGCGCGCC 426
426 AGATCGTTGGCGGAGTATACCTTGTGGCGCGCAGGGGCCCCAGGTTGGGTGGTGGCGCA 485
427 AGATCGTTGGTGGAGTTTACTTGTTCGCGCGCAGGGGCCCCAGTGGGTGGTGGCGCA 486
486 CAAGGAAGACTTTCGAGCGGTCCAGCCACGTTGGAAGCGCCAGCCCATCCCTAAAGATC 545
487 CGAGAAAGACTTTCGAGCGGTCCAGCCACCTCGAGGTAGAGTCAAGCTATCCCAAGGCTC 546
546 GCGCGCTCCACTGCGCAATCTTGGGGAAACACAGGATACCCCTGGGCCCTTATACGGAATG 605
547 GTGGCCCGAGGCGAGGACTTGGGCTCAGCCCGGTACCTTGGGCCCTTATGCAATG 606
606 AGGAGCTCGGCTGGGAGGATGGTCTGTTCGCGCGAGGTTCCCGTCTCTTGGGGCC 665
607 AGGGCTGCGGGTGGGCGGATGGTCTGTCTCTCCCGTGGCTCTCGGCGCTAGTGGGGCC 666
666 CCAATGACCCCGGCATAGTTCGCGCAAGTGGGTAAAGTCAATCATACCTTAAAGTGG 725
667 CCACAGACCCCGGCGTAGGTTCGCGCAATTTGGGTAAAGTCAATCATACCTTACGTCG 726
726 GCTTTGCCGACCTCATGGGGTACATCCCTGTGCGGGCGCCCGCTCGCGGCGTCCGCA 785
727 GTTTCGCGACCTCATGGGGTACATACCGCTGTGCGGGCCCTCTTGGAGGCGTCCCA 786
786 GAGCTCTCGCGATGCGGTGAGAGTCTTGGAGGAGCGGGTTAAATTTGCAACAGGGAATC 845
787 GGGCCCTGGCGCATGGCGTCCGGGTTCTGGAAGAGCGCGTGAATATGCAACAGGGAATC 846
846 TACCGGTGCTCTTTCTATCTTCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 905
847 TTTCTGTTGCTCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 906
906 TCTCGGCTGCCGAAGTGAAGAACATCATGATCCGCGCTACATGGTGAATACGACTGACCA 965
907 CTTTCGGCTTACCAAGTGGCGCACTCCACGGGGCTTTTACCACTGATGATGTCCTTA 966
966 ATGACAGCATTACCTGGAGCTCCAGGCTGCTGCTCTCCAGCTCCCGGGTGGTCCCGT 1025
967 ACTCGAGTATTGTATACGAGGCGCGCATGCCATCTGTCACACTCCGGGGTGGTCCCTT 1026
1026 GCGAGAAAGTGGGAATGATCTCTAGTGTGATACCGGTCTCACCGAATGTGGCGGTGC 1085
1027 GCGTTCGTGAGGCGAACCGCTCGAGGTGTGGGTGGCGATGACCCCTACGTTGGCCACCA 1086
1086 AGCGGGCGGGCGCTTACGAGGGGTTCGCGAGCGACATCGACATGTTGTGATGTCGG 1145
1087 GGGATGGCAAACTCCCGCGAGCGAGCTTCGAGCTCACATCGATCTGCTTGTTCGGAGCG 1146
1146 CACAGCTCTGCTGCTGCTCTAGTGGGGGAGCTCTGCGGTGGGTGATGCTCGAGGCC 1205
1147 CCACCTCTGTTGCGGCGCTTCTAGTGGGGGACCTATGCGGGTCTGTCTTCTTGTTCGGCC 1206
1206 AATGTTTCACTGCTGCGCGAGCACCTGTTTGTCCAAGACTGCAATGCTTCTCTCT 1265
1207 AACTGTTTCACTTCTTCTCCAGGCGCACTGGAACGCAAGTTGCAATGCTTCTCTCT 1266
1266 ACCCTGTTACCATCACTGACACCCGCAATGGCATGATGATGATGATGATGATGATGATGAT 1325

```

2335	DB	2335	TACTGCTGACCACTACACAGTGGCAGGTCCTCCCGTGTTCCTTCAACACCTACACAGCCT	2394
2406	QY	2406	TGTCGACTGGTCTCTCCACCTCCACAAAACATCGTGGACGTACAATTCATGTATGGCC	2465
2395	DB	2395	TGTCACCCGGCCTCATCCACCTCCACACAGAACATTTGTGACGTGCAGTACTTTGTACGGGG	2454
2466	QY	2466	TATCACCTGCCCTCAAAATATACATCGTCCGATGGGAGTGGGTAAATCATCTTTATTCCTGC	2525
2455	DB	2455	TGGGGTCAAGCATCCGCTCTGGGCCATTAAGTGGGAGTACGTCTCTCTCTGTTCCTTTC	2514
2526	QY	2526	TCCTTAGCGAGCCGACAGGGTTTGGCGCTGCTTATATGATGCTCATCTTTGTGGCCAGGCCG	2585
2515	DB	2515	TGCTTGCAGACCGCGCGGTCTGCTCTCTGCTGTGGATGATGCTACTCATATCCAAAGCGG	2574
2586	QY	2586	AAGCAGCACTAGAGAAGCTGGTCACTTTGCACGCTCGCAGCGCAGCTAGCTGCAATGGCT	2645
2575	DB	2575	AGCGCGCTTTGGAGAACCTCGTAACTAATGACAGCATCCCTGGCCGGACGCACGGTCT	2634
2646	QY	2646	TCCTATATTTTGTCTATCTTTTTCGTGGCTGCTGTTGTGTACATCAAGGGTGGGTAGTCCCT	2705
2635	DB	2635	TTGTATCTCTCTCGTGTCTTCTGCTTTGCAATGGTATTTGAAGGGTAAGTGGGTGCCG	2694
2706	QY	2706	TAGCTACCTATTTCCTCACTGCGCTGTGCTCTTTAGCCTACTGTCTTAGCATTTGCCCC	2765
2695	DB	2695	GAGCGGTCTACACCTTTTAGCGGATGTGCGCTCTCTCTGCTCTCTGTTGGCGTTGCCCC	2754
2766	QY	2766	AACAGCGTTATGCTTTATACGCATCTGTGCATGGCCAGATAGGAGCGGCTCTGCTGGTAA	2825
2755	DB	2755	AGCGGGCGTACCGCTGGACAGGAGTGGCCGCTCGTGTGGCGGTGTGTTCTCGTCG	2814
2826	QY	2826	TGATCACTCTCTTTACTCTCACCCCGGGTATAAGACCTTCTCAGCCGGTTTGTGTGT	2885
2815	DB	2815	GGTTGATGGCGCTGACTCTGTCAACCATATTAACAAGCGCTATATCAGCTGGTGTGTGT	2874
2886	QY	2886	GTTTGTGCTATCTTCAGCCCTGGGGGAAGCTATGGTCCAGAGTGGGCACCACTATGC	2945
2875	DB	2875	GGCTTCAGTATTTTCTGACCAAGTGGGAAGCGCACTGCACGTGTGGAATTCCTCCCTCA	2934
2946	QY	2946	AGGTGCGGGTGGCGGTGATGGCATCATATGGGCGCTCGCCATATTTACCCAGGTGTGG	3005
2935	DB	2935	ACGTCCGAGGGGGCGCGAOGCGTCACTTACTCATGTGTCTGTACACCCGACTCTGG	2994
3006	QY	3006	TGTTTGAATATAACCAAGTGGTCTTTGGCGGTGCTTTGGGCTTCTTACCTCTTAAAGTGG	3065
2995	DB	2995	TATTTGACATCACCAAAATGTGCTGGCGCTCTTCGGACCCCTTTGGATTCTTCAAGCCA	3054
3066	QY	3066	CTTTTACGCGGTGCGGTACTTCTGCTCAGGCTCAGCTCTACTAGGAGTGCACCATGG	3125
3055	DB	3055	GTTTCTTTAAAGTACCTTACTTTGTGCGCGTCAAGGCGCTCTCGGTCTCGCGGTTAG	3114
3126	QY	3126	CAAGGCATCTCGCGGGGCGAGGTACGTCCAGATGGCGCTACTAGCCCTTTGGCAGGTGA	3185
3115	DB	3115	CGCGAAGATGATCGAGGGCCATTACGTGCAAAATGGTCATCATTAAGTTAGGGGCGTTA	3174
3186	QY	3186	CTGGCACTTACATCTATGACCACTCAACCCCTATGTGGAATGGGCTGCTAGTGGCCTGC	3245
3175	DB	3175	CTGGCACCTATGTTTATAACCATCTCACTCTCTTTCGGGACTGGCGCACCAACGGCTTC	3234
3246	QY	3246	GGGACCTGGCGGTGCGGTTGAGCCCTCATCTTCAGTCCGATGGAGAAGTCAATTG	3305
3235	DB	3235	GAGATCTTGGCGGTGGCTGTAGAGCCAGTCTGCTTCTCCAAATGGAGACCAAGCTCATCA	3294
3306	QY	3306	TCCTGGGAGCGGAGACAGCTGCTGTGGGGACATTTTACAGCGACTTCCCGTGTCCGCC	3365
3295	DB	3295	CGTGGGGGCGATACCGCGCGGTGCGGTGACATCATCAACGGCTTGCTGTTCGCGCC	3354
3366	QY	3366	GACTTGTGGAGGTCTCTCTTGGCCCGCAGCTGATGGCTTATACCTCCAAAGGGGTGGAGTC	3425
3355	DB	3355	GCAGGGGCGGGAGATACTGCTCTGGCGCAGCCGATGGAAATGCTCTCCAAGGGGTGGAGT	3414
3426	QY	3426	TTCTCGCCCCCATCTGCTTACGCCGACGACAGAGTGGCCCTTTTGGGCACCATATGTTG	3485
3415	DB	3415	TGCTGGCGCCATCACGGGCTACGCCGACGACAGAAGGGGCCCTCTAGGGTGCATTAATCA	3474

QY	3486	TGAGCATGACGGGCGGACAAAGACAGACAGAGCTGGGGAAATTCAGGTCTCTCCACAG	3545	QY	4656	CGGCGCTTGGGGTATGGGCTTGAACCTCAGTGGGCACTACTACAGAGGTTGGACGTCTCCG	4625
DB	3475	CGAGCTTAATGCGCGGGAACAAACCAAGTGAAGGTGAGGTCCAGATTGTGTCACTG	3534	DB	4555	CAAGCTGTGTCGCAATTTGGGCACTCAATGCGGTGCGCTACTACCGGCTCTTGACGTCTCCG	4614
QY	3546	TCATCTAGTCTTCTCGGAACATCCATCTCGGGGTTTTGTGGAGTGTCTACCATGGAG	3605	QY	4626	TAATACCAACTCAGGAGACGTAGTGTGTCGCCACCGACCCCTCATGACAGGTTATA	4685
DB	3535	CTGCCAAACCTTCTCGGAACGTGCATCAATGGGGTGTCTGAGTGTCTACCAAGGG	3594	DB	4615	TCATCCGACCAAGCGGATGTTGTGCTGGGCAACCGATGCCCTCATGACCGGCTATA	4674
QY	3606	CTGGCAACAAGACTCTGGCGGCTCAGCGGGTCCGGTCAGCAGATGTACTCAGTGTCTG	3665	QY	4686	CTGGGAGCTTTGACTCTCGTGTGCTCAACAGTAGCGGTCACTCAAGTTGTAGACTTCA	4745
DB	3595	CCGGAACAGGAGCCATCGCTCAACCAAGGGTCTGTCTATCCAGATGTATACCAATGTAG	3654	DB	4675	CGGGGACTTTGACTCTCGGTGATGACTGCAATAGTGTGTCAACCGACAGCTGATTTCA	4734
QY	3666	AGGGGACTTTAGTGGGTGCGGACAGCCCTCGGACTTAATCTTTGGAGCGGTGCACTG	3725	QY	4746	GTTTAGACCCCAATTTACCATTAACCAACACAGATTGTCCCTCAAGACGTGTCTACGTA	4805
DB	3655	ACCMAGACTTTGGGTGCGGCTCGGCTCGCAAGTAGCCGCTCATTTGACACCTGACATT	3714	DB	4735	GCCTTGACCTTACCTTACCAATTGAGACAATCACGCTCCCGAGGATGCTGTCTCCGCA	4794
QY	3726	GTGAGCGGTGCGACTGTACTGTGTGTCAGCGGACGCTGATGTCTATCCCGGCTCGAAGAC	3785	QY	4806	GCCAGCGCGGGTGCACGGGTAGGGAAGACTGGGCAATTTATAGTATGTTTCCACTG	4865
DB	3715	CGGCTCTCGGACCTTTACTGTCTACGAGGACGCGGATGTCAATTCCTACTTTGAAAGGCTCT	3774	DB	4795	CTCAAGCTCGGGGAGGACTGGCAGGGGAAACGAGCATCTACAGATTTGTGGCACCGG	4854
QY	3786	CGGGGACAAACGGGAGCGGTACTCTCCCGGAGACCTCTTCCACTTGAAGGGTCTCT	3845	QY	4866	GTGAGGAGCTCAGGATGTTTGTGACGTGTGTCTGTGAGTGTCTACGACGACGAGGG	4925
DB	3775	GGGGTGATAGCAGGGGAGCGTGTGTGCGCCCGGCGCATTTCTCTACTTTGAAAGGCTCT	3834	DB	4855	GGGAGCGCCCTCCGCGATGTTGACTGTGCTGTGAGTGTCTCTGTGAGTGTCTATGACGAGGT	4914
QY	3846	CAGGAGCGCGGTATGCGCCAGGGGCAACGCTGTGCGAGTCTTCCGGGACGCTGTGT	3905	QY	4926	CGGCTATGATGAGCTCACACATCGGAGACACCGTCAAGGCTCAGGGCGTATTTCAACA	4985
DB	3835	CGGGGGTGTGCTGTGTGCGCCCGGGGACGCGGTGGGCAATTTAGGGGCGCGGTGT	3894	DB	4915	GTGCTTGTATGAGCTCACGCGCGGAGACTACAGTTAGGCTACGAGGCTACATGAACA	4974
QY	3906	GCTCTCGGGGCGTGGCTAAGTCAATAGATTTCATCCCGGTTGAGACACTCGACATCGTCA	3965	QY	4986	CGCCCGGTTTGTGTGTGCGCAAGACCATCTTGTAGTGTGGAGGAGTGTTCACCGGCC	5045
DB	3895	GCACCGTGGAGTGGCTAAGGCGGTGACTTTATCCCTGTGGAGAACTTAGAGACAACA	3954	DB	4975	CGCCCGGCTTCCGCTGTGCGGAGGACCATTTTGAATTTTGGGAGGCGCTTTTACAGGCC	5034
QY	3966	CGCGTCCCGCACCTTTAGTGAACAAGACACACACCTGTGTGCGCCAGACCTATCAGG	4025	QY	5046	TCACACATAGATGCCACTTCTTCCCAACAAAGCAATCGGGGAAATTTTCGCAAT	5105
DB	3955	TGAGTCCCGGCTTACAGGATTAATCTCTCCACAGTAGTGCCTCCAGAGCTTCCAGG	4014	DB	5035	TCATCATATAGATGCCACTTCTATCCAGAAAGAGAGTGGGAGAACTTCTCTT	5094
QY	4026	TCGGGTACTTGCATGCGCCGACTGGAGTGGAAAGACACCAAGTCTCTGTGCGCATATG	4085	QY	5106	ACTTAACAGCTACAGGCTACAGGTGTGCGCTAGGCGCAAGGCCCCCGGCTCTGGG	5165
DB	4015	TGGCTACCTCCATGCTCCACAGGAGCGGCAAGACACAGGTCCCGGCTGCAATG	4074	DB	5095	ACCTGTAGGTTACCAAGCCACCGTGTGCGCTAGGCTCAAGCCCTCCCGCATCTGGG	5154
QY	4086	CTGCTCAGGGGTATAAGTGTCTGTCTTAACTCCCTCAGTGGCTGCGCACCTCGGGTTG	4145	QY	5166	AGTCTATGTGGAAGTGTGTGACTCGACTCAAGCCCACTCTGGGGCCCCACCTCTCC	5225
DB	4075	CAGCTCAGGGGTATAAGGTGTCTACTTCAACCCCTCTGTGTGTCACAACTGGGCTTG	4134	DB	5155	ACCAGATGTGGAAGTGTGTGATTTCGCTCAAGCCCACTCTCCATGGGCCCAACCCCTGC	5214
QY	4146	GGGGTACTTGTCTAAGGACATGGCATCAATCCCAACATTTAGGACTGGAGTCAGACTG	4205	QY	5226	TGTACCGCTTGGGCTCTGTTTACCAAGGCTCACCTCACATCCCGTGACGCAATAACA	5285
DB	4135	GTGCTTACATGTCCAAAGGCTCATGGGATCGATCTTAACATCAGGACCGGGGTGAGAACAA	4194	DB	5215	TATACAGACTGGGCGTGTGTGAGATGAATCACTCTGACGACCCAGTCAACCAATACA	5274
QY	4206	TGACGACGGGGCGCCATCAGTACTCCACATATGGCAAAATTCCTCGCGATGGGGCT	4265	QY	5286	TGCCACCTGCAATGCAAGCGGCTTGTAGGTCAATGACAGCAGATGGGTCTTGGCAGGG	5345
DB	4195	TTACACTGGAGCCCATCAGTACTCCACCTACGGCAAGTTCCTTGGCGACGGCGGT	4254	DB	5275	TCATGACATGATGTGCGCGGACCTGGAGGTCTGACAGGACCTGGGTGCTCTGGCG	5334
QY	4266	GTGCGGGCGGCGCTACGACATCATATGTGATGAATGCCATGCGCGTGGACTTACCA	4325	QY	5346	GAGTCTTGGGGCGGTGCGCGGTATGTGCTGGGACCGGGTGTGTTCATCATCGGCC	5405
DB	4255	GCTCGGGGGGCTTATGACATAATATTTGTGACGAGTGCCACTCCACGGATGCCACAT	4314	DB	5335	GGGTCTGCTGTTTGGCGCGGTATGCTGTCTCAACAGGCTGCGGTGCTCATGTGGGCA	5394
QY	4326	CCATCTTGGCATCGGAACAGTCTCTGATCAAGCAGAGACAGTGGGGTCAGACTTAATG	4385	QY	5406	GCTTGCACATTAACAGCAGCGCTGTGTCGCGCGGACAAAGAGGTCCTCTATGAGGCTT	5465
DB	4315	CCATCTTGGCATCGGCACTGTCTTGACCAAGCAGAGACTGGCGGGGCGAGACTGTTG	4374	DB	5395	GGGTCTGCTTGTTCGGGAAAGCGGCAATCATCTGACAGGGAAGTCTCTTACCGAGGT	5454
QY	4386	TGCTGGCTACAGCTACGCGCCCTGGGTGAGTCAACCCCGACCCCAACATAGAGGAG	4445	QY	5466	TTGATGAGATGGAGGAATGTGCTCTTAGGGCGGCTCTCATTTGAAGAGGGGCGAGTAG	5525
DB	4375	TGCTGCGCACCGCCACCCCTCGGGCTCGGTCACTGTGCGCCCATCCCAACATCGAGGG	4434	DB	5455	TCGATGAGATGGAAGTGTCTCTCAGCACTTACCGTACATCGAGCAAGGATGATGCTCG	5514
QY	4446	TGGGCTTGGGAGGAGGCGAGATCCCTTTCTATGGGAGGGCGAATTCCTCTGTCTTACA	4505	QY	5526	CCGAGATGTGAGTCCAAAGATCCAAAGGTATTGAGCAGAGGTTTCCAAACAGCTCAAG	5585
DB	4435	TTGCTCTGTCCACACCGAGAGATCCCTTTTACGCAAGGCTATCCCTCTCGAAGTAA	4494	DB	5515	CCGAGAGTTCAGCAGAAAGGCGCTCGGCTCTCTGACAGCCGCTCCCTCGAGGAGAGG	5574
QY	4506	TCAAGGAGGAGACATCTGATCTTTGCGCATTTCAAGAAAGAAAGTGTGACAGCTCGCG	4565	QY	5586	ACATACAACCCCTGTGCGAGGCTTCTATGGCCCAAGGTAGAACTTCTGGGCCAAACACA	5645
DB	4495	TCAAGGGGGGAGACATCTCATCTTCTGTCTATTTCAAGAAAGAGTGGCAGCACTCGCG	4554	DB	5575	TTATCGCCCTGTGTCAGACCACTGGCAAAACTCGAGACCTTCTGGGCGAGGATA	5634
				QY	5646	TGTGGAACCTTCTATGCGGCATCCAAATACCTCGAGGACTTATCAACACTGCCAGGAAAC	5705

||||| 5635 TGTGGAACCTTTCATCAGTGGGATCAATATCTTGGGGCTTGTCAACGCTGCTGGTAACC 5694 Db
||||| 5706 CTGAGTAGCTTCCATGATGGCGTTAGTGGCGCCCTCACCAGTCCGCTGTCACAAAGCA 5765 Qy
||||| 5695 CCGCCATTTGCTTCAATGATGGCTTTTACAGCTCTGTCAACAGCCCACTAACACATAGCC 5754 Db
||||| 5766 CCACTATCTTCTCAACATTTTGGGGGGCTGGCTAGCATCCCAAATTTGACACACCGCGG 5825 Qy
||||| 5755 AAACCTCTCTTCAACATATTGGGGGGTGGTGGCTGCGCAGCTCGCGCCCCGGTG 5814 Db
5826 GGGCCACTGCTTGTGTTGTCAGTGGCTAGTGGAGTGCCTGAGCAGTATAGGCTTAG 5885 Qy
||||| 5815 CCGCTACTGCTTGTGGCGCTGGCTTAGCTGGCGCGCATCGGCAGTGTGGACTGG 5874 Db
5886 GTAAGGTCTAGTGGACATCTGCGAGGATGATGGTGGGGCATTTGGGGGCTCTGTCG 5945 Qy
5875 GGAAGTCTCATAGACATCTTTCAGAGGTATGGCGGGCTGGGGGAGCTCTTGTGG 5934 Db
5946 CATTCAGATCATGCTTGGCGAGAACCCCTCCATGGAGGATGTCGTAACCTTCTGCTGCTG 6005 Qy
5935 CATTCAGATCATGAGCGGTGAGGTCCCTCCACGGAGGACTGGTCACTACTACTGCCCCG 5994 Db
6006 GAATTTCTTCCGGGTGCTTGGTAGTGGAGTCACTGCGCGGCCATTTCTGCGCGGAC 6065 Qy
5995 CCATCTCTCGCGGAGCCCTGCTAGTGGCGTGGTCTGTGCAAGCAATACTGCGCGGC 6054 Db
6066 ACGTGGACCGGGGAGCGCGCTCAATGGATGATAGACTCATTTGCTTCTTCCCA 6125 Qy
6055 ACGTGGCCCGGGGAGGGGAGTGCAGTGGATGAACCGGCTGATAGCCCTTGGCTCCC 6114 Db
6126 GAGGAATACAGCTCGCCGCCACCACTACGTCAGGAGTGGAGTGGCTGCACGCTGTGA 6185 Qy
6115 GGGGGAACCATTTTCCCCACGCACTACGTCGGGAGGAGTGCAGCTGCGCGGTCA 6174 Db
6186 CCCAACTACTTGGTCCCTTACCAATACCAAGCTGCTCAGAGACTCCAACTGATTA 6245 Qy
6175 CTGCCATCTCAGCAGCTCACTGTAACTAACAGCTCCTGAGCGGACTGCACAGTGGATA 6234 Db
6246 CTGAGGACTGCCCATCCATCGCGGGCTGCTGGCTCCGCGATGTTGGGACTGGTTT 6305 Qy
6235 GCTCGGAGTACCACTCCATGCTCGGTTCTGGCTTAAGGGACATCTGGGACTGGATAT 6294 Db
6306 GCACATCTTAAACAGCTTTAAATAATGGCTGACCTCCAAATTAATCCCAAAGATGCCG 6365 Qy
6295 GCGAGGTGTGAGGACTTTAAGACTGGCTTAAAGCTAAGCTCATGCCACAGCTGCTG 6354 Db
6366 GCCTCCCTTTGTCTCTGTCAAAAGGGGTACAAGGGCGTGTGGCGGCACTGGCATCA 6425 Qy
6355 GGATCCCTTTGTCTCTGCGCAGCGGGGTATAAGGGGGTCTGGCGAGTGGACGGCATCA 6414 Db
6426 TGACCACAGGTGCTTGGCGGCCAATATCTTGGCAATGTCGCTTGGGCTCCATGA 6485 Qy
6415 TGCACATCTGCTGCACTGTGGAGCTGAGATCACTGGACATGTCAAAAACGGGACGATGA 6474 Db
6486 GAATCAGGGGCTTAAGACCTGATGAATATCTGGCAGGGGACTTTCTTATCAATTGTT 6545 Qy
6475 GGATCGGTCTTAGACCTCGAGGAACATGTGGAGTGGGACCTTCCCATTTAATGCTT 6534 Db
6546 ACACGGAGGCGAGTGGCTGCGGAAACCCCGGCCAACTTTAAGGTGGCCATCTGAGGG 6605 Qy
6535 ACACCGGGCCCTGTACCCCTCTCTGCGCGGAACTACAGTTTCGCGCTATGGAGG 6594 Db
6606 TGGCGGCTCAGAGTACCGGAGGTGACGAGCAGCAGGTCATACCACTACATAACAGGAC 6665 Qy
6595 TGTCTGAGAGGAATATGTGGAGATAAGGAGGTGGGGGACTTCCACTAGTGAACGGGTA 6654 Db
6666 TCACCACTGATACTCAAGAGTCCCTGCAACTTACCTCTCCCGAGTCTCTTTTCTGGG 6725 Qy
6655 TGACTACTGACATCTCAATGCCCCGTGCCAGGTCCCATCGCCGGAATTTTTCACAGAT 6714 Db
6726 TGGAGGAGTGCAGATCCATAGTTTGGCCCCCAGCACCGAAGCGGTTTTCGCGGATGAGG 6785 Qy
||||| 6715 TGGACGGGGTGGCCCTACATAGTTTGGCGCCCCCTGCAAGCCCTTGTCTGCGGGAGGAGG 6774 Db
6786 TCTCGTCTGCTGGCTTGGCTTAAATTTGCTGCTGGGTCCAGCTTCTCTGCGGACCCCTG 6845 Qy
6775 TATCATTTAGAGTAGACTTCCAGANATACCCGGTAGGGTGCANATTTACCTTCCGAGCCCG 6834 Db
6846 AACCGGACACAGACGCTATTGATGTCATGCTAACAGATCCATCTCATATACCGGGGAGA 6905 Qy
6835 AACCGGACGTCGCGCTGTGAGCTCCATGCTCACTGATCCCTCCATATAACAGCAGAGG 6894 Db
6906 CTGACGCGCGGCTTTAGCGCGGGGTCAACCCCATTCGAGGCAAGCTCTCTGCGGAGCC 6965 Qy
6895 CGCGCGGGCGAAGTTGGCGAGGGATCAACCCCTCTGTGGCCAGCTCTCTGCGCTAGCC 6954 Db
6966 AGCTATCGGACCATCTGCTGCGAGCACCTGCACACACACCGCAAGCTATGATGTCG 7025 Qy
6955 AGCTATCGGCTCCATCTCTCAAGGCAACTTGCACGCTTAACCATGACTCCCTGATGCTG 7014 Db
7026 ACATGCTGATGCTTAACCT-----GTTTCATGGGGGGCGATGTGACTCGGATAG 7073 Qy
7015 AGCTCATAGAGGCCAACCTCTATGAGGCGAGAGATGGCGGCAACATCACAGGGTTG 7074 Db
7074 AGTCTGGGTCCAAAGTGTGTTCTGAGCTCTCTCGAACCCTTGTGGAATCTGTGGAATGCTG 7133 Qy
7075 AGTCAGAAAAACAAAGTGTGATTTCTGAGCTCTCTCGATCCGCTTGTGGCGGAGGAGCG 7134 Db
7134 ACCTTGGGCTTGCATACCATCAGATATACATGCTCCCCAAGAAAGGTTTCCACACAGCTT 7193 Qy
7135 AGCGGAGATCTCCGTACCCGCGAAGATCTCTCGGAGTCTCGGAGATTTGCCCCAGGCC 7194 Db
7194 TACCGGCTTGGGACGCGCTGATTAACAACCCCTTGTGGAATCTGTGGAATGCTGGAATGCTG 7253 Qy
7195 TGCGGCTTGGCGCGCGCGGACTATAACCCCTGCTAGTGAGAGCTGTGGAAGAGCCG 7254 Db
7254 ATTACGACCGGCTGCTGTTGGGGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7313 Qy
7255 ACTACGACCACTCTGTGTGCTCAATGCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7314 Db
7314 CTCCCCAAGGAGCGCGGACAGTGGGCTTAAGTGAGGACTCCATAGGAGATGCCCTTC 7373 Qy
7315 CTCCGCTCGGAGAGCGGAGCGGTGCTCTCACTGAATCAACCTATCTACTGCTTGG 7374 Db
7374 AACAGTGGCCATTAAGTCTTTTGGCGAGCCCCCCCCCAAGCGGCGATTCAGGCTTTTCA 7433 Qy
7375 CCGAGCTCGCCACAGAGACTTTGGAGCTCTCAACTTCCGCGCATTAACGGCGCAATA 7434 Db
7434 CGGGGCGGGCTGCTGCGGATTCGCGAGTCAAGCGCTCTCTCTCTCTCTCTCTCTCTCTCT 7493 Qy
7435 CGACCAACATCTCTG-----AGCGCGGCTCTCTGCTGCCCCCGGACTCGG 7482 Db
7494 AGACAGGTTCCATCTTTCATGCCCCCTCGAGGGGAGCTTCGAGATCCAGACCTGG 7553 Qy
7483 ACGTGAAGTCTTCTCTCTCATGCCCCCTCGAGGGGAGGCTCGGAGATCCGATCTTA 7542 Db
7554 AGCCTGAGCAGGTAGAGCCCCCAACCCCCCGGGGGGTGGCAGCTCCCGGCTCGG 7613 Qy
7543 -----GGAGG 7548 Db
7614 ACTCGGGTCTGGTCTACTTCTCGAGGAGGAGCATCTCGTCTGTGCTGCTCATGT 7673 Qy
7549 GGTCTATGTTCAACGGTCAAGTAGTGAGGCGCAACCGGAGGAGTGTCTGTGCTCAATGT 7608 Db
7674 CATACTCTGGACCGGGCTCTAATACTCTTGTAGTCCCGAAGAGGAGATTAACGA 7733 Qy
7609 CTTACTCTTGGACAGGCGCTCTGTCACCCCTGCGCGGAGAGACAGAACTGCCCA 7668 Db
7734 TTAACCCCTTGGAGCAACTCTCTTGTGCGATATCAACAAGGTGTACTGTATCAACAACA 7793 Qy
7669 TCAATGACTAAGCAACTCTGTTGCTACGTCACCAAAATTTGTTGTTATTTCCACCACTAC 7728 Db
7794 AGAGCGCTCACTAAGGGCTAAAAAGGTAACTTTTGTATAGGATCAAGTGTCTGACTCT 7853 Qy
7729 GCAGTGTCTTCCAAAGGAGAGAAAGTCACTTTTGCACAGACTGCAAGTCTTCTGGACGCG 7788 Db

8934 TGCTCTGATGACACACCTTTCTTCCATTTCTCATGGCTCAAGACAGCTGGAGCCAGAAC 8993
 8869 TGATACTGATGACCCATTTCTTTAGCGTCTTTATAGCCAGGACAGCTTTGAACAGGCC 8928
 8994 TCAACTTTGAGATGTACGAGCGGTGTACTCCGTGAGTCCCTTGGACCTCCAGCTATTA 9053
 8929 TCGATTGCGAGATCTACGGGGCTGTCTACTCCATGAACCACTTGATCTACCTCCAATCA 8988
 9054 TTGAAAGGTTACATGGGCTTGACGCTTTTCTCTGCACACATACACTCCCCCAAGACTCA 9113
 8989 TTCAAAGACTCCATGGCTTCAGCGCATTTTCACTCCACAGTTACTCTCCAGGTGAATTA 9048
 9114 CACGGGTGGCTTCAGCGCTTCAGAAACTTTGGGGCGCCACCTCTCAGAGCTGGAGAGCC 9173
 9049 ATAGGGTGGCGCATCGCTCAGAAACTTTGGGGTACCGCCCTTGGAGCTTGGAGACCC 9108
 9174 GGGCAGCTGACGTACAGGGGCTCCCTCATCTCCCGTGGGGGAGAGCGGCGGTGTGGGTC 9233
 9109 GGGCCCGAGAGCGTCCCGCTTAGGCTTTGGCCAGAGAGGAGGCGTGCCTATGTGGCA 9168
 9234 GATATCTCTTCAATTTGGCGGTGAAGACCAAGCTCAAACTCACTCCATTTGCCGAAGCGC 9293
 9169 AGTACCTCTTCAACTGGGCGGTGAAGACCAAGCTCAAACTCACTCCATTTAGCGCGCTG 9228
 9294 GCCTCTCTGATTTATCCAGCTGGTTTACCGTCCGGCGCGCGGGGCGAGCATTTATCACA 9353
 9229 GCCAGCTGGACTTTGTCGGCTGGTTTACCGCTGGCTTACAGCGGGGAGACATTTATCACA 9288
 9354 GGTGTGCGGTGCGCGACCCCGCTTATGCTCTTTGGGCTACTCTCTACTTTTGTAGGG 9413
 9289 GGTGTGCTCATGCGCGCGCGCTGGATCTGGTTTGGCTTACTCTCTGCTTGTGCGAGGG 9348
 9414 TAGGCTTTTCTTACTTCCCGCTCGGTAGAG 9444
 9349 TAGGCATCTACCTCTCTCCCAACCGATGAAG 9379

Search completed: February 12, 2004, 22:41:28
 Job time : 639 secs

7854 ACTACGACTCAGTCTTAAGGACATTAAGCTAGCGGCTCCAAAGGTCAACGCAAGGCTCC 7913
 7789 ATTACAGGACGTAATCAAGAGGTTAAAGCAGCGGCTCAAAAGTGAAGCTTAAGTTC 7848
 7914 TCACCTGAGAGGCTTGCAGTTAACCCACCCCATTTCTGCAAGATCTAAATATGGGT 7973
 7849 TATCCGTAGAGGAGCTTGCAGCTGACGCCCCACACATCAGCCAAATCCAAGTTGGTT 7908
 7974 TTGGGGTAAAGAGGCTCCGAGCTTGTCCGGGAGGCGGTAAACCATCAAGTCCGTGT 8033
 7909 ATGGGCAAAAGACGTCGTTGTCATGCCAGAAAGCGGTAAACCCACATCAACTCCGTGT 7968
 8034 GGAAGGCTCTTGGAGACTCAGAAACCAATTTCCCAACCATTTATGGCCCAAAATG 8093
 7969 GGAAGGCTCTTGGAGAAAGCAATGTAAACCAATAGACATACCATCATGGCTTAAGAACG 8028
 8094 AGGTGTTCTGCTGACCCCAACCAAGGGGGGCAAGAGCAGCTCGCTTTATCGTTTACC 8153
 8029 AGGTTTCTGCTTACGCTTGAGAGGGGCTCGTAAGCCAGCTCGTCTCATCGTTTC 8088
 8154 CTGACCTCGGCTCAGGCTTGGAGAGATGGCCCTTTATGACATTAACAAAACTTC 8213
 8089 CCGATCTGGGCGTGGCGGTGCGAAAGATGGCTTTGTACGAGCTGGTTTACAAAGCTCC 8148
 8214 CTCAGGCGGTGATGCGGGCTTCTTATGATTTCCAGTATTTCCCGCTTCAGCGGTAGAGT 8273
 8149 CTTTGGCGGTATGGAGCTCTTACGATTTCCATCTACCAAGACAGCGGTGAAT 8208
 8274 TTTCTTTGAAGCATGGGCGGAAAGAGGACCCCTATGGGTTTTTCTGATGATACCGAT 8333
 8209 TCTCGTGAAGCGTGAAGTCCAAAGAAACCCCAATGGGTTCTCGTATGATACCGCT 8268
 8334 GCTTTGATCAACCGTCACTGAGAGAGACATCAGGACTGAGAGTCCATATATCGGCGCT 8393
 8269 GCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACGAGAGGCAATCTACCAATGTT 8328
 8394 GCTCTTTCCGAGGAGGCCACACTGCCATACACTGCTTAATCTGAGAGACTTTAGTGG 8453
 8329 GTGACCTCGACCCCAAGCGCGGTGCGCCATCAAGTCCCTCACGAGAGGCTTTATGTTG 8388
 8454 GAGGCTTATGTTCAACAGCAAGGCGCAAACTCTCGGGTACAGCGGTTGCGCGCCAGCG 8513
 8389 GGGCCCTCTTACCAATTCAAGGGGGAGAACTGCGGCTATCGCAGGTGCGCGCGAGCG 8448
 8514 GGTGTCTCACCATAGCATGAGGAAACCATCATGCTAGTGAAGAGCTTACGCGCTT 8573
 8449 GCGTACTGACAACTAGTGTGTGTAACACCCCTCACTTGTCTACATCAAGGCGCGGCGAGCT 8508
 8574 GTAAAGCTGACGGGATAATCGCGCCCAATGCTGTGTATGCGGATGACTTGTGTTCA 8633
 8509 GTCAGCGCGAGGCTCTCAGGACTGACCATCTGCTGTGTGCGGAGCATTTAGTCTGTTA 8568
 8634 TCTCAGAAAGCCAGGAGCCAGGAGGAGCAGCGGAACTCTGAGAGCTTCAACGAGGCTA 8693
 8569 TCTGTGAAGCGCGGGGTCCAGGAGGAGCGCGGAGCTTGAAGCTTCAACGAGGCTA 8628
 8694 TGACAGGATATCTGCGGCTCTGCTGAGACCCCGCCAGACCGGAGTATGATCTGGAGCTGA 8753
 8629 TGACCAAGTACTCGCGGCTCTGCGGAGACCCCGCCCAACCAAGATACGACTTGGAGCTCA 8688
 8754 TAACTCTGCTCTCAAAATGTTGTTCTGCGGCTGGGCCCAACAGCGCGCGGAGATACT 8813
 8689 TAACTATGCTCTTCAACGTTGTCAGTCCGCCACAGAGCGGCTGGAAGAGGCTTACT 8748
 8814 ACCTGACCAAGAGCCCTACCACTCCAAATCGCCCGGCTGCTGGGAAACAGTTAGACACT 8873
 8749 ACCTCACCGGTGACCCCTACAAACCCCTCGGAGAGCTGCGTGGGAGACAGCAACA 8808
 8874 CCCCTGCAATTTAGCTGGGAAACATATCCAGTAGCGCCCGCCCATATGGGCTCGCA 8933
 8809 CTCCAGTCAATTTCTGGCTAGGCAACAATAATCATGTTTGGCCCCCACACTGTGGGCGAGGA 8868

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 12, 2004, 06:01:01 ; Search time 17628 Seconds
(without alignments)
13388.970 Million cell updates/sec

Title: US-09-980-559-1
Perfect score: 9711
Sequence: 1 accgcgcctaataggcg.....ggctctctgcagatcatgt 9711

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 1215238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

- 1: em_estba.*
- 2: em_esthum.*
- 3: em_estin.*
- 4: em_estnu.*
- 5: em_estov.*
- 6: em_estpl.*
- 7: em_estro.*
- 8: em_hc.*
- 9: gb_est1.*
- 10: gb_est2.*
- 11: gb_hc.*
- 12: gb_est3.*
- 13: gb_est4.*
- 14: gb_est5.*
- 15: em_estfun.*
- 16: em_estom.*
- 17: em_gss_hum.*
- 18: em_gss_inv.*
- 19: em_gss_pln.*
- 20: em_gss_vrt.*
- 21: em_gss_fun.*
- 22: em_gss_mam.*
- 23: em_gss_mus.*
- 24: em_gss_pro.*
- 25: em_gss_rod.*
- 26: em_gss_pbg.*
- 27: em_gss_vrl.*
- 28: gb_gss1.*
- 29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	274.6	2.8	488	9	AV755731
c 2	235.8	2.4	492	9	AV758366
c 3	120.2	1.2	1201	13	BX386383
c 4	118.6	1.2	891	13	BU843859

5	117.8	1.2	915	13	BX371184
6	117.4	1.2	980	12	BM415785
7	117.4	1.2	1201	13	BX425171
8	117.2	1.2	1101	29	CNS0080M
9	117.2	1.2	1201	13	BX446581
10	116.8	1.2	809	14	CD521141
11	116.6	1.2	338	13	BX403712
12	116.4	1.2	716	10	BF348329
13	116.4	1.2	1110	9	AL515173
14	116.2	1.2	336	9	AL730440
15	116.2	1.2	801	9	AV758179
16	116.2	1.2	981	13	BX349784
17	116.2	1.2	732	13	BX447068
18	116.1	1.2	871	13	BU589926
19	116.1	1.2	965	13	BO921874
20	116.1	1.2	1043	10	BF791952
21	115.8	1.2	233	13	BX425119
22	115.8	1.2	403	14	CB768447
23	115.8	1.2	428	13	BX425361
24	115.8	1.2	964	13	BU962355
25	115.6	1.2	279	13	BX425800
26	115.6	1.2	370	13	BX425084
27	115.6	1.2	446	13	BX403933
28	115.4	1.2	590	14	CB937830
29	115.4	1.2	1073	13	BQ930284
30	115.2	1.2	597	9	AV762619
31	115.2	1.2	889	13	BX439710
32	115.2	1.2	907	13	BU936114
33	115.2	1.2	924	14	CA970305
34	115.2	1.2	974	13	BX428246
35	115.1	1.2	489	12	BI377034
36	115.1	1.2	889	9	AL666715
37	115.1	1.2	896	29	CNS008P8
38	115.1	1.2	914	13	BX432884
39	115.1	1.2	916	13	BX390165
40	115.1	1.2	945	13	BX390676
41	115.1	1.2	972	13	BU531322
42	115.1	1.2	1201	13	BX403767
43	115.1	1.2	1238	10	BG179896
44	114.8	1.2	528	12	BG910163
45	114.8	1.2	592	14	CA377114

ALIGNMENTS

RESULT 1
AV755731/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES

AV755731 488 bp mRNA linear EST 19-OCT-2000
AV755731 BM Homo sapiens cDNA clone BMFAKB03 5', mRNA sequence.
AV755731.1 GI:10913579
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,
Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng,
L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G.,
Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and Chen, Z.
Homo sapiens cDNA BM clones
Unpublished
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

Location/Qualifiers

Copied from 09540843 on 05/19/2004

Copied from 09540843 on 05/19/2004

9177	CACGTGCAGTCAGGGCGTCCCTCATCTCCGGTGGGGGAGAGCGCGGTTTCGGTTCGAT	9236
	:: :	
792	CMCVKGTAAAAAMAMCMCWTTGTTKCMYTKRVCNCAACCCVVBKKGCMCKTTTKGACARY	733
	:: :	
9237	ATCTCTTCAATTGGCGGTGAAGACCAAGCTCAAACTCACTCCATTGCCGGAAGCGCGCC	9296
	:: :	
732	AYTGADRMCMHTTGTTKKRARMCMCMCSMGSKMSMCAACGBKGTKKKGABTKMC	673
	:: :	
9297	TCCTGGATTTATCAGCTGGTTTCAACGTGGCGCCGGGGGGCGACATTTATCAACGC	9356
	:: :	
672	KCCKGVMGVAVBKMCSVBKGCACCCCKGCCGRMCKGRTKGRKKGVCSYBTGTGGSGAG	613
	:: :	
9357	TGTCCGCTGCCGACCCCGCTATTGCTCTT-----TGGCCCTACTCTCTACTTTTGG	9407
	:: :	

[illegible]

```

RESULT 9
LOCUS      BX446581
DEFINITION BX446581 1201 bp mRNA linear EST 22-MAY-2003
             Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
             CLOBB002ZG04 5-PRIME, mRNA sequence.
ACCESSION  BX446581
VERSION     BX446581.1
KEYWORDS    GI:31021737
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1201)
             Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
             Full-length cDNA libraries and normalization
             Unpublished
CONTACT     Genoscope

```

Genoscope - Centre National de Sequencage
BP 131 91006 EVRY cedex - France
Email: secre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : liang@lifetech.com URL :
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CLOB8002G04RP1.
Location/Qualifiers
1. 1201

```

BASE COUNT
ORIGIN
286 a 182 c 159 g 300 t 274 others
/obj_type="mRNA"
/db_xref="taxon:9606"
/clone="CLO8B002ZG04"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/note="vector: pcMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pcMVSPORT 6 vector.
Library was not normalized."

```

BASE COUNT

382 a	150 c	141 q	136 t
-------	-------	-------	-------

CA). Note: this is a NIH_MGC Library."

	BASE COUNT	382 a	150 c	141 g	136 t	
	ORIGIN					
	Query Match	1.2%	Score 116.8;	DB 14;	Length 809;	
	Best Local Similarity	75.5%;	Pred. No. 0.13;			
	Matches 145; Conservative	0;	Mismatches 47;	Indels 0;	Gaps 0;	
	9438	GGTAGAGCGGCACACATTAGTCTACTCCTCATGCTAACTGTGCCCTTTT				9497
	Qy					TTTTTTTTTTTTTTTTTTTT
	Db	249	GGGGGGGGGGCCCCCAATTTT			TTTTTTTTTTTTTTTTTTTT
						190
	9498	TT				9557
	Qy					TTTTTTTTTTTTTTTTTTTT
	Db	199	TT			130
						TTTTTTTTTTTTTTTTTTTT
	9558	TTTTTTTTTCTTTTTTCTTTTCCTTCTTTTCTTACTTAACTTTCTTTCCCTGGTG				9617
	Qy					TTTTTTTTTTTTTTTTTTTT
	Db	129	TT			70
						TTTTTTTTTTTTTTTTTTTT
	9618	GCTCATCTTAG	9629			
	Qy					
	Db	69	GGTTCATTGTAG	58		

RESULT 11	338 bp	linear	EST 13-MAY-2003
BX403712		mRNA	
LOCUS			
DEFINITION	BX403712 Homo sapiens PLACENTA Homo sapiens CDNA clone CLOBA009ZE09		
ACCESSION	3-PRIME, mRNA sequence.		
VERSION	BX403712		
KEYWORDS	BX403712.1 GI:30638967		
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 338)		
JOURNAL	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.		
COMMENT	Full-length cDNA libraries and normalization Unpublished Contact: Genoscope Genoscope - Centre National de Sequencage Bp 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by life technologies, a division of Invitrogen. This sequence belongs to sequence cluster 1321.r For more information about this cluster, see http://www.genoscope.cns.fr/ http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CLOBA009ZE09FF1&cluster=1321.r Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Paradise Avenue Genoscope sequence ID : CLOBA009ZE09FF1.		

```

FEATURES
  source
    Location/Qualifiers
      1..338
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="CLOBA009ZE09"
        /tissue_type="PLACENTA"
        /clone_lib="Homo sapiens PLACENTA"
        /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
        with a NotI-oligo(dT) primer. Five prime end enriched,
        double-strand cDNA was digested with Not I and cloned
        into the Not I and EcoRV sites of the pCMVSPORT 6 vector.
        Library was not normalized."
      60 a      42 c      38 g      32 others
      ORIGIN
        Query Match      1.2%; Score 116.6; DB 13; Length 338;
        Best Local Similarity 73.8%; Pred. No. 0.23;
        Matches 144: Conservative 3; Mismatches 48; Indels 0; Gaps 0;

```

Query Match	1.2%	Score 116.6;	DB 13;	Length 338;
Best Local Similarity	73.8%;	Pred. NO. 0.23;		
Matches 144:	Conservative	3;	Mismatches 48;	Indels 0;
	Capable	0;		

Query Match	1.2%	Score 116.6;	DB 13;	Length 338;
Best Local Similarity	73.8%;	Pred. NO. 0.23;		
Matches 144:	Conservative	3;	Mismatches 48;	Indels 0;
	Capable	0;		

Qy 9439 GTAGAGCGGCACATTAGCTACCTCCTAGCTACCTAGCTACCTGTCCTCTTTT 9498
 Db |||||
 49 GTAAATTTGGGCCCCCTCGAGGATACCTCTAGAGCGCGCCCTTTT 108
 Qy 9499 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9558
 Db |||||
 109 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 168
 Qy 9559 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9618
 Db |||||
 169 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 228
 Qy 9619 CTCACCTTACCCCT 9633
 Db |||||
 229 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT

RESULT 12
 BF348329/c
 LOCUS
 DEFINITION 716 bp mRNA linear EST 22-NOV-2000
 602022359f1 NCI CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4157534
 5', mRNA sequence.
 ACCESSION BF348329
 VERSION BF348329.1 GI:11295913
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 716)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Straubeberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: David N. Louis, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1AM9432 row: d column: 15
 High quality sequence stop: 316.

FEATURES
 source
 1..716
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4157534"
 /tissue_type="anaplastic oligodendroglioma with lp/19q
 loss"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NCI CGAP Brn67"
 /note="Organ: Brain; Vector: pCMV-SPORT6; Site 1: NotI;
 site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 2.3 kb. Constructed by Life
 Technologies. Note: this is a NCI CGAP Library."
 BASE COUNT 452 a 89 c 126 g 49 t
 ORIGIN
 Query Match 1.2%; Score 116.4; DB 10; Length 716;
 Best Local Similarity 69.1%; Pred. No. 0.16;
 Matches 159; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

Qy 9381 TGCTCTTTGGCCTACTCTACTTTTGTAGGGGTAGCCCTTTCTACTCCCGCTCGGT 9440
 Db |||||
 392 TTCTCTTTGTCTTCTGTCCTGTTGTTGGGCGGTTTCCGTTTCTCCCGCTCGGT 333
 Qy 9441 AGACGGGCACATTAGTACACATCCATAGTACTGTCCTCTTTTCTTTTCTTTT 9500
 Db |||||
 332 CGGTTTTCTCCCTCCCGCCCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 273

Qy 9501 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9560
 Db |||||
 272 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 213
 Qy 9561 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9610
 Db |||||
 212 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 163
 RESULT 13
 AL515173
 LOCUS
 DEFINITION 1110 bp mRNA linear EST 08-MAY-2003
 CL0BB028ZH02 3-PRIME, mRNA sequence.
 ACCESSION AL515173
 VERSION AL515173.2 GI:30465058
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1110)
 AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 COMMENT On Feb 13, 2001 this sequence version replaced gi:12778666.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 9050.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CL0BB028ZH02FPI&cluster=9050.r. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Paraday Avenue Genoscope sequence ID : CL0BB028ZH02PPI.
 FEATURES
 Location/Qualifiers
 1..1110
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CL0BB028ZH02"
 /tissue_type="NEUROBLASTOMA"
 /clone_lib="Homo sapiens NEUROBLASTOMA"
 /note="Vector: pCMVSPORT6; 1st strand cDNA was primed
 with a NotI-oligo(dt) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."
 BASE COUNT 169 a 309 c 98 g 324 t 210 others
 ORIGIN
 Query Match 1.2%; Score 116.4; DB 9; Length 1110;
 Best Local Similarity 67.0%; Pred. No. 0.12;
 Matches 142; Conservative 4; Mismatches 56; Indels 0; Gaps 0;
 Qy 9399 TACTTTTGTAGGGGTAGGCTTTTCTCTACTCCCGCTCGGTAGGGGCACATTAGC 9458
 Db |||||
 12 TDNNNTTTNTNNNNNNNNNNNNNTASTTTTNGGCGCCCTCGAGGATCTCTAGAGC 71
 Qy 9459 TACACTCCATAGCTAATCTGTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9518
 Db |||||
 72 GCGCGCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 131
 Qy 9519 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9578
 Db |||||
 132 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 191
 Qy 9579 TTCCTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9610
 Db |||||
 192 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 223

RESULT 14	
AL730440/c	
LOCUS	AL730440 linear EST 18-APR-2002
DEFINITION	Danio rerio embryonic inner ear subtracted cDNA Danio rerio cDNA clone BNOAA113ZH08 5', mRNA sequence.
ACCESSION	AL730440
VERSION	AL730440.1
KEYWORDS	GI:20195044
SOURCE	EST.
ORGANISM	Danio rerio (zebrafish)
	Danio rerio
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
REFERENCE	1 (bases 1 to 336)
AUTHORS	Coimbra,R., Weild,D., Brottier,P., Blanchard,S., Levi,M., Hardelin,J.P., Weissenbach,J. and Petit,C.
TITLE	A subtracted cDNA library from the zebrafish (Danio rerio) embryonic inner ear
JOURNAL	Unpublished
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

[illegible][illegible]

Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and Chen, Z.

Homo sapiens cDNA EM clones

Unpublished

Contact: Zequang Han
Chinese National Human Genome Center at Shanghai
3351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China

Tel: 86-21-50801919(ex. 45)
Fax: 86-21-50801922

Email: hanzg@chgc.sh.cn

This clone is available at CHGC in Shanghai.

FEATURES

```

source
1008601
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMPAFW09"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/note=vector; pTriplex2; Site 1: sfIIA; Site 2: sfIIB"
469 a 96 c 108 q 136 t 2 others
BASE COUNT
```

[illegible]

Search completed: February 12, 2004, 12:18:15
Job time : 17634 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 11, 2004, 23:09:14 ; Search time 22809 Seconds

(without alignments)
5439.910 Million cell updates/sec

Title: US-09-980-559-2

Perfect score: 16120

Sequence: 1 MSTNPKPQKTKENTRRPQ.....LLFGLLLLVGVGLFLPAR 3033

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n_model -DEV=xlh
-Q-/cgm2 1/USPRO_spool/US09980559/runat_09022004_141953_26401/app_query.fasta_1.3207
-DB=GenEmbl -QMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USPR=US09980559 @CGN 1 16399 @runat_09022004_141953_26401 -NCFU=6 -ICPU=3
-NO MAP -LARGEQUERY -NGG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pin.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sv.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	16120	100.0	9711	6	AX057086 Sequence
2	16120	100.0	9711	6	AX057317 Sequence
3	16120	100.0	9711	6	AX057395 Sequence
4	16120	100.0	9711	14	AF177036 Sequence
5	15810	98.2	9589	14	HPCPOLP
6	15824	98.2	9589	6	E07361
7	15824	98.2	9589	6	E07362
8	15824	98.2	9589	6	E07362
9	15426	95.7	9700	14	AF169005
10	15394	95.3	9659	14	AB047641
11	15356	95.3	9416	14	AF238483
12	15355	95.3	9416	14	AF238482
13	15354	95.2	9674	14	AB047640
14	15350	95.2	9691	14	AB047644
15	15347	95.2	9416	14	AF238481
16	15324	95.1	9683	14	AB047645
17	15322	95.0	9416	14	AF238484
18	15321	95.0	9416	14	AF238485
19	15311	95.0	9661	14	AF169002
20	15266	94.7	9653	14	AF169004
21	15249	94.6	9693	14	AF169003
22	15207	94.3	9678	14	AB047642
23	15115	93.8	9641	14	AB047643
24	14943	92.7	9678	6	BD160781
25	14943	92.7	9678	14	AB047639
26	14356	89.1	9513	14	D50409
27	14185	88.0	9488	14	AB031663
28	13971	86.7	9511	14	HPCJ8G
29	13885	86.1	9654	14	AB030907
30	13884	86.1	9416	14	AF238486
31	13810	85.7	9511	6	E07468
32	13810	85.7	9511	6	I12865
33	13803	85.6	9511	6	E07467
34	13049	80.9	9611	6	AX057088
35	13049	80.9	9611	6	AX057092
36	13049	80.9	9611	14	AF177037
37	13049	80.9	9611	14	AF177039
38	12925	80.2	9611	6	AX057090
39	12925	80.2	9611	6	AX057094
40	12925	80.2	9611	14	AF177038
41	12925	80.2	9611	14	AF177040
42	12032	74.6	9379	14	AF165058
43	12031	74.6	9605	6	AX739971
44	12031	74.6	9605	14	HCJ238799
45	12028	74.6	9379	14	AF165060

ALIGNMENTS

RESULT 1

AX057086
 LOCUS AX057086 9711 bp DNA linear PAT 17-JAN-2001
 DEFINITION Sequence 1 from Patent WO0075338.
 ACCESSION AX057086
 VERSION AX057086.1 GI:12309919
 KEYWORDS Hepatitis C virus
 SOURCE Hepatitis C virus
 ORGANISM Hepatitis C virus
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
 1
 REFERENCE Yanagi, M., Bukh, J., Emerson, S.U. and Purcell, R.H.
 AUTHORS Cloned genome of infectious hepatitis C virus of genotype 2a and
 TITLE uses thereof
 JOURNAL Patent: WO 0075338-A 1 14-DEC-2000;
 THE GOVERNMENT OF THE UNITED STATES OF AMERICA (US)
 FEATURES
 source 1. 9711
 /organism="Hepatitis C virus"
 /mol_type="genomic DNA"
 /db_xref="taxon:11103"
 BASE COUNT 1987 a 2853 c 2668 g 2203 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 9711
 Score: 16120.00 Matches: 3033
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-980-559-2 (1-3033) x AX057086 (1-9711)
 QY 1 MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgProGln 20
 DB 341 ATGAGCACAAATCTAAACCTCAAGAAACCAAGAAACCAACCAACCTGCGCCACAA 400
 QY 21 AspValLysPheProGlyGlyGlyGlnIleValGlyValTyrLeuLeuProArgArg 40
 DB 401 GACGTTAAGTTTCCGGCGCGCGCCAGATCGTTGGCGGAGTATACITTTGTCGCGCAGG 460
 QY 41 GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly 60
 DB 461 GCCCCAGGTTGGTGTGCGCGCGCAAGAGAGACTTCGAGCGGTCCAGCCACGTTGA 520
 QY 61 ArgArgGlnProLysAspArgSerThrGlyLysSerTrpGlyLysProGly 80
 DB 521 AGCGCCAGCCCATCCCTAAAGATCGCGCTCCACTGGCAAAATCCTGGGAAACACAGGA 580
 QY 81 TyrProTrpProLeuTyrGlyAsnGluGlyLeuGlyTrpAlaGlyTrpLeuLeuSerPro 100
 DB 581 TACCCCTGGCCCTTATACGGGAATGAGGAGACTCGGCTGGGCGAGGATGGCTCTGTCCCC 640
 QY 101 ArgGlySerArgProSerTrpGlyProAsnAspProArgHisArgSerArgAsnValGly 120
 DB 641 CGAGGTTCCGCTCCCTCTTGGGGGCCCAATGACCCCGGCGATAGTTCGCGCAAGTGGGT 700
 QY 121 LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProValVal 140
 DB 701 AAGGTCATCGATACCCTAACGTGCGGCTTTCCCGACCTCATGCGGTATACATCCCTGTGCTG 760
 QY 141 GlyAlaProLeuGlyGlyValAlaArgAlaLeuAlaHisGlyValArgValLeuGluAsp 160
 DB 761 GCGCGCCGCTCGCGCGCGTCCGAGAGCTCTCGCGCATGCGGTGAGAGTCTCTGGAGGAC 820
 QY 161 GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla 180
 DB 821 GGGGTTAAATTTGCAACAGGAACTTACCCGGTGTCTCTCTTTCTATCTTCTTCTGCTGCC 880
 QY 181 LeuLeuSerCysIleThrThrProValSerAlaAlaGluValLysAsnIleSerThrGly 200
 DB 881 CTGCTGTCTGCATCACCAACCCCGGTCTCCGCTGCGAAGTGAAGAACATCAGTACCGGC 940

QY 201 TyrMetValThrAsnAspCysThrAsnAspSerIleThrTrpGlnLeuGlnAlaVal 220
 DB 941 TACATGGTGACTAACGACTGACCAATGACAGCATTACCTGGCAGCTCCAGGCTGCTGC 1000
 QY 221 LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTrpIle 240
 DB 1001 CTCACGCTCCCGGGTGGTCCGTCGAGAAAGTGGGAATGTCATCTCAGTCTGGATA 1060
 QY 241 ProValSerProAsnValAlaValGlnArgProGlyAlaLeuThrClnGlyLeuArgThr 260
 DB 1061 CCGGTCTCCACCAATGTGCGGTGACGCGCGCGGCGCTCACGAGGCTTTCGCGACG 1120
 QY 261 HisIleAspMetValValMetSerAlaThrLeuCysSerAlaLeuTyrValGlyAspLeu 280
 DB 1121 CACATCGACATGGTGTGATGTCGCGCACGCTCTGCTCTGCTCTACCTGGGGACCTC 1180
 QY 281 CysGlyGlyValMetLeuAlaAlaGlnMetPheIleValSerProGlnHisTrpPhe 300
 DB 1181 TCGGTGGGTGATGCTCGCAGCCCAATGTTTCAATGTCTCGCGCAGCACCACCTGGTTT 1240
 QY 301 ValGlnAspCysAsnCysSerIleTyrProGlyThrIleThrGlyHisArgMetAlaTrp 320
 DB 1241 GTCCAAAGACTGCAATTTGCTTCCATCTACCTGGTACCATCTGACACCCGATGGCATGG 1300
 QY 321 AspMetMetMetAsnTrpSerProThrAlaThrMetIleLeuAlaTyrAlaMetArgVal 340
 DB 1301 GACATGATGATGAACCTGTCGCGCAGGCTACCATGATCTTGGCGTACCGGATGCTGTC 1360
 QY 341 ProGluValIleIleAspIleIleSerGlyAlaHisTrpGlyValMetPheGlyLeuAla 360
 DB 1361 CCGGAGGTCAATTATACATCATATTAGCGGGCTCATTTGGGCGTCTATGTTGGCTTGGCC 1420
 QY 361 TyrPheSerMetGlnGlyAlaTrpAlaLysValValIleLeuLeuAlaGly 380
 DB 1421 TACTTCTCTATGAGGAGCGTGGCGCAAGTCTGTTCATCTCTGTTGGCGCGCGG 1480
 QY 381 ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr 400
 DB 1481 GTGACGCGCGCACCATCTACTGTTGGGGTTCGCGCGCAGACACCGCGGCGCTCACC 1540
 QY 401 SerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTrp 420
 DB 1541 AGCTTATTTGACATGGCGCCAGGCAAAATCCAGTCTGTAAACCAATGGCAGCTGG 1600
 QY 421 HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer 440
 DB 1601 CACATCAACGCGCACCCGCTGAACATGCAATGACTCTTGCACACCCGCTTTATCGGCTCT 1660
 QY 441 LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg 460
 DB 1661 CTGTTCTACACCCACAGCTTCAACTCGTCAGGATGTCGCGAAGCATGTCGCTGCGCG 1720
 QY 461 SerIleGluAlaPheArgValGlyTrpGlyAlaLeuGlnTrpGluAspAsnValThrAsn 480
 DB 1721 AGTATCGAGGCTTCCGGTGGGATGGGCGCTTGCATATGAGGATAATGTCAACCAAT 1780
 QY 481 ProGluAspMetArgProTyrCysTrpHisTyrProProArgGlnCysGlyValSer 500
 DB 1781 CCAGAGGATATGAGACCTTATTCGTGGCACTACCCACCAAGGCGAGTGTGGGTGCTCC 1840
 QY 501 AlaLysThrValCysGlyProValTyrCysPheThrProSerProValValValGlyThr 520
 DB 1841 GCGAAGACTGTGTGGCCGAGTACTGTTTACCCCGACCGGAGTGTAGTGGGACG 1900
 QY 521 ThrAspArgLeuGlyAlaProThrThrTrpTrpGlyGluAsnGluThrAspValPheLeu 540
 DB 1901 ACCGACAGGCTTGGAGCGCCCACTTACGCTGGGGGAGAAATGAGACAGATGTCTTCTTA 1960
 QY 541 LeuAsnSerThrArgProProLeuGlySerTrpPheGlyCysThrTrpMetAsnSerSer 560
 DB 1961 TTGAACAGCACTCGACACCGCTGGGGTCTATGGTTCGGCTGCACGTGGATGAATCTTCT 2020

Qy	561	GlyTyrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer	580	Db	3101	GCTCTACTGAGTGTGCACCATGGCAAGGCATCTCGCGGGGGCAGGTACGTCCAGATG	3160
Db	2021	GGCTACACCAAGACTTGGCGCGCACCACTCGCGTACTAGAGCTGACTTCAACGCCAGC	2080	Qy	941	AlaLeuLeuAlaLeuGlyArgTrpThrGlyThrTyrIleTyrAspHisLeuThrProMet	960
Qy	581	ThrAspLeuLeuCysProThrAspCysPheArgLysHisProAspThrThrTyrLeuLys	600	Db	3161	GGCTACTAGCTTGGCAGGTGGACTGGCATTATCTATGACCACTTCACCCCTATG	3220
Db	2081	ACGGACCTGTGTGGCCCAACGGACTGTGTAGGAAGCATCTCTATACCACTTACCTCAA	2140	Qy	961	SerAspTrpAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIleLeuPhe	980
Qy	601	CysGlySerGlyProTrpLeuThrProArgCysLeuIleAspTyrProTyrArgLeuTrp	620	Db	3221	TCGGATTGGCTGCTAGTGGCTCGGGACCTGGCGGTGCGGTGAGCTTATCATCTTC	3280
Db	2141	TGGGGCTCTGGGCCCTGCTCACGCCAAGGTGCTGATCGACTACCCCTACAGCTCTGG	2200	Qy	981	SerProMetGluLysLysValIleValTrpGlyAlaGluThrAlaAlaCysGlyAspIle	1000
Qy	621	HisTyrProCysThrValAsnTyrThrIlePheLysIleArgMetTyrValGlyGlyVal	640	Db	3281	AGTCCGATGAGAAAGTCAATGTCTGGGAGCGGAGACAGCTGCTTGTGGGGACATT	3340
Db	2201	CAATTACCCCTGCACAGTTAACTATACCATCTTCAAAATAAGGATGTATGTGGAGGGTT	2260	Qy	1001	LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp	1020
Qy	641	GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp	660	Db	3341	TTACACGGACTTCCCGGTGTCGCCGACTTGGTGGGAGGTCTCTTGGCCGAGCTGAT	3400
Db	2261	GAGCACAGGCTCACGGCTGCATGCAATTTCACTCGTGGGATCGTTGCAACTTGGAGGAC	2320	Qy	1021	GlyTyrThrSerLysGlyTrpSerLeuLeuAlaProIleThrAlaTyrAlaGlnGlnThr	1040
Qy	661	ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrGluTrpAlaIleLeuPro	680	Db	3401	GGCTATACCTCCAAAGGGTGGAGTCTTCTCGCCCCCATCTGCTTACGCCACAGACA	3460
Db	2321	AGAGACAGAGTCACTGTCTCTTGTGTCATCTCCACACGGAATGGGCCATTTTACCT	2380	Qy	1041	ArgGlyLeuLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGluGlnAla	1060
Qy	681	CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisLeuHisGlnAsnIle	700	Db	3461	CGTGGCTTTTGGGCACCATAGTGTGAGCATGACGGGGCGGACAAAGACAGACAGGCT	3520
Db	2381	TGCTCTTACTCGGACCTCGCCGCTTGTGACTGGTCTCTCCACCTCCACCAAAACATC	2440	Qy	1061	GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly	1080
Qy	701	ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrLysTyrIleValArgTrp	720	Db	3521	GGGGAAATTCAGGTCTGCTGCCACAGTCACTAGTCTCTCTCGGAACATCCATCTCGGG	3580
Db	2441	GTGGACGTACAAATTCATGTATGGCTATACCTGCTCCCTCAAAATATACGTCCGATGG	2500	Qy	1081	ValLeuTrpThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro	1100
Qy	721	GluTrpValIleLeuLeuPheLeuLeuAlaAspAlaArgValCysAlaCysLeuTrp	740	Db	3581	GTITTTGTGACTGTCTACCATGGAGCTGGCAACAGACTCTGGCCGGCTCACGGGTCCG	3640
Db	2501	GAGTGGGTAACTCTTATTTCTGCTCTTAGCGACCCGAGGTTCGCGCTGCTTATGG	2560	Qy	1101	ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTrpProSerProGly	1120
Qy	741	MetLeuIleLeuLeuGlyGlnAlaGluAlaLeuGluLysLeuValIleLeuHisAla	760	Db	3641	GTCCAGCAGATGTACTCCAGTGTCTAGGGGGACTTAGTAGGGTGGCCAGCCCTCTGGG	3700
Db	2561	ATGCTCATCTTGTGGCCAGCGCGAGCAGCAGTACAGAAAGCTGGTCACTTTGCAGCT	2620	Qy	1121	ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArgAsn	1140
Qy	761	AlaSerAlaAlaSerCysAsnGlyPheLeuTyrPheValIlePhePheValAlaAlaTrp	780	Db	3701	ACTAAATCTTTGGAGCGGTGCACGTGTGGAGCGGTGACCTGTACTGTGTGCAGCGGAAC	3760
Db	2621	GCAGCGCAGCTAGCTGCAATGGCTCTCTATATTTTGTCTATCTTTTGTGGTGTCTTG	2680	Qy	1141	AlaAspValIleProAlaArgArgArgGlyAspLysArgGlyAlaAlaLeuLeuSerProArg	1160
Qy	781	TyrIleLysGlyArgValValProLeuAlaThrTyrSerLeuThrGlyLeuTrpSerPhe	800	Db	3761	GCTGATGTCTATCCCGCTCGAAGACCGGGGGACAAACGGGGAGCGCTACTCTCCCGAGA	3820
Db	2681	TACATCAAGGTCGGTAGTCCCTTAGCTACCTATTTCCTCACTGGCCCTGTGTCTTT	2740	Qy	1161	ProLeuSerThrLeuLysGlySerSerGlyGlyProValLeuCysProArgGlyHisAla	1180
Qy	801	SerLeuLeuLeuAlaLeuProGlnAlaTyrAlaTyrAspAlaSerValHisGly	820	Db	3821	CCTCTTTCCACCTTGAAGGGGTCTCAGGAGGCCCGGTGCTATGCCCCAGGGGCCACGCT	3880
Db	2741	AGCCTACTGCTCTAGCATTTGCCCAACAGGCTTATGCTTATGAGGCATCTGTGCAATGGC	2800	Qy	1181	ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle	1200
Qy	821	GlnIleGlyAlaAlaLeuLeuValMetIleThrLeuPheThrLeuThrProGlyTyrLys	840	Db	3881	GTCCGAGTCTTCGGGCAGCTGTGTGCTCTCGGGGGGTGGCTAAGTTCATAGATTTCATC	3940
Db	2801	CAGATAGAGCGGCTCTGCTGGTAAATGATCACTCTCTTTTACTCTCAACCCCGGGTATAAG	2860	Qy	1201	ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro	1220
Qy	841	ThrLeuLeuSerArgPheLeuTrpTrpLeuCysTyrLeuLeuThrLeuGlyGluAlaMet	860	Db	3941	CCCGTTGAGACACTCGACATCGTCAGCGGTCCCCCACCTTTAGTGCACCAACAGCACCA	4000
Db	2861	ACCTCTCTACCGGGTTTTTGTGGTGTGTGTCTATCTCTGACCCCTGGGGGAGCTATG	2920	Qy	1221	ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLys	1240
Qy	861	ValGlnGluTrpAlaProProMetGlnValArgGlyGlyArgAspGlyIleIleTrpAla	880	Db	4001	CCTGCTGTCCCGACACCTATCAGGTGGGTACTTGCATGTGCCCGAGTGGCAGTGAAG	4060
Db	2921	GTCCAGAGTGGGCACCATATGACAGTGGCGGTGGCGGTGATGGCATCATATGGGCC	2980	Qy	1241	SerThrLysValProValAlaTyrAlaGlnGlyTyrLysValLeuValLeuAsnPro	1260
Qy	881	ValAlaIlePheTyrProGlyValValPheAspIleThrLysTrpLeuLeuAlaValLeu	900	Db	4061	AGCACAAAGTTCCTGTCTGCATATGCTCGAGGGGTATAAAGTGTAGTGTCTTAATCCC	4120
Db	2981	GTGCCATATTTACCCAGGTGTGGTGTGTGACATAACCAAGTGGCTCTTGGCGGTGCTT	3040	Qy	1261	SerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerLysAlaHisGlyIleAsnPro	1280
Qy	901	GlyProAlaTyrLeuLeuLysGlyAlaLeuThrArgValProTyrPheValArgAlaHis	920	Db	4121	TCAGTGGCTGCCACCTCTGGGGTTTGGGGGTCTTGTCTAAGGCATGTGCATCAATCCC	4180
Db	3041	GGGCTGCTTACCTCTCAAAAGGTGCTTTGACGGCGGTGCGTACTCTGTCAGGCGCTCAC	3100	Qy	1281	AsnIleArgThrGlyValArgThrValThrThrGlyAlaProIleThrTyrSerThrTyr	1300
Qy	921	AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyGlyArgTyrValGlnMet	940				

4181	AAACATTAGAGCTGGAGTACAGGACTGTGACGACCGGGGGCCCCATCAGCTACTCCACATAT	4240
1301	GlyLysPheLeuAlaaspGlyGlyCysAlaGlyGlyAlaTyrAspIleIleIleCysAsp	1320
4241	GGCAAAATTCCTCGCGATGGGGCTGTGGGGCGGGCTACGACATCATATGTGAT	4300
1321	GluCysHisAlaValaspSerThrThrLeuGlyIleGlyThrValLeuaspGlnAla	1340
4301	GAATGCCATCGCGTGAGCTCTACCAACCATCTCTGGCATCGGAACAGTCTCTTGATCAAGCA	4360
1341	GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProGlySerValThr	1360
4361	GAGACAGCTGGGGTCAGACTAACTGTGTGGCTACAGCTACGCCCCCTGGGTGAGTGACA	4420
1361	ThrProHisProAsnIleGluGluValAlaLeuGlyGlnGluGlyGluIleProPheTyr	1380
4421	ACCCCCCACCACATAGAGAGGTGGCCCTTGGCAGGAGGGCGAGATCCCTTCTTAT	4480
1381	GlyArgAlaIleProLeuSerTyrIleLysGlyIleArgHisLeuIlePheCysHisSer	1400
4481	GGGAGGGCATTCCTCTGTCTACATCAAGGAGGAAGACATCTGATCTTCTGCCATTCA	4540
1401	LysLysLysCysaspGluLeuAlaAlaAlaLeuArgGlyMetGlyLeuAsnSerValAla	1420
4541	AAGAAAAAGTGTGACAGCTCGCGCGGCCCTTTCGGGGTATGGGTGTAACCTCAGTGCCA	4600
1421	TyrTyrArgGlyLeuaspValSerValIleProThrGlnGlyAspValValValAla	1440
4601	TACTACAGAGGGTTGACAGCTCTCCGTATATCCACTCAGGGAGAGCGTAGTGGTCTGCC	4660
1441	ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnVal	1460
4661	ACCGAGCGCCTCATGACGGGTATACTGGGGACTTTGACTCCGTGATCGACTGCCAACGTA	4720
1461	AlaValThrGlnValValaspPheSerLeuAspProThrPheThrIleThrGlnIle	1480
4721	GGGGTCACTCAAGTTGTAGACTTCAGTTTTAGACCCCCACATTCACCATAACACACAGATT	4780
1481	ValProGlnAspAlaValSerArgSerGlnArgGlyArgThrGlyArgGlyArgLeu	1500
4781	GTCCCTCAAGACGCTGTCTCAGTAGCCAGCGCCGGGTGCGACGGGTAGGGGAGAGCTG	4840
1501	GlyIleTyrArgTyrValSerThrGlyGluArgAlaSerGlyMetPheAspSerValVal	1520
4841	GGCAITTTATAGGTATGTTCTCCACTGCTGAGCGAGCCTCAGGAATGTTTCACAGTGTAGTG	4900
1521	LeuCysGluCysTyrAspAlaGlyAlaAlaTyrTyrGlyLeuLeuThrProSerGluThrThr	1540
4901	CTCTGTGAGTGTACACGCGAGGGCGGATGTGTATGAGCTCACACATCGAGAGCACCC	4960
1541	ValArgLeuArgAlaTyrPheAsnThrProGlyLeuProValCysGlnAspHisLeuGlu	1560
4961	GTCAAGCTCAGGGCGATTTTCAACACGCCGGTTTGCCTGTGTGCCAAGACCATCTTGAG	5020
1561	PheTrpGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThr	1580
5021	TTTTGGGAGGCAGTTTTTACCGGCCCTCACACATAGATGCCACATCTCTTTCCCAACA	5080
1581	LysGlnSerGlyLysAsnPheAlaTyrLeuThrAlaTyrGlnAlaThrValCysAlaArg	1600
5081	AAGCAATCGGGGAAAAATTTTCGATATCTTAAACAGCCCTACAGGGCTACAGTGTGGCTAGG	5140
1601	AlaLysAlaProProSerTrpAspValMetTrpLysCysLeuThrArgLeuLysPro	1620
5141	GCCAAAGCCCCCCCCCGTCTCGGAGCGTATGTGGAAAGTTTGTGACTCGACTCAAGCCC	5200
1621	ThrLeuValGlyProThrProLeuLeuTyrArgLeuGlySerValThrAsnGluValThr	1640
5201	ACACTCGTGGCCCCACACTCTCTCTGTACCGCTTGGGCTCTGTATTACCAACGAGGTCA	5260
1641	LeuThrHisProValThrLysTyrIleAlaThrCysMetGlnAlaAspLeuGluValMet	1660
5261	CTCACATCCCGTGAAGAAATACATCGGCACCTCGATGCAAGCCGACCTTGAGGTCA	5320

1661	ThrSerThrTrpValLeuAlaGlyGlyValLeuAlaValAlaAlaTyrCysLeuAla	1680
5321	ACCAGCACATGGGCTTTGGCAGGGGGAAGTCTTGGCGCGCTGGCGCGTATTGGCTGGCG	5380
1681	ThrGlyCysValCysIleIleGlyArgLeuHisIleAsnGlnArgAlaValValAlaPro	1700
5381	ACCGGGTGTGTTTGCATCATCGGCGCTTGACATTTAACACGAGCGCGCTGTTGGCGCG	5440
1701	AspLysGluValLeuTyrGluAlaPheAspGluMetGluGluCysAlaSerArgAlaAla	1720
5441	GACAAGGAGGTCTCTATTAGGCTTTTCATGAGATGGAGGAATGTGCCTCTAGGCGCGCT	5500
1721	LeuIleGluGluGlyGlnArgIleAlaGluMetLeuLysSerLysIleGlnGlyLeuLeu	1740
5501	CTCATTTGAAGAGGGGCGACGGATACCGAGATGCTGAAGTCCAAGATCCAAGGCTTATTG	5560
1741	GlnGlnAlaSerLysGlnAlaGlnAspIleGlnProThrValGlnAlaSerTrpProLys	1760
5561	CAGCAGGCTTCCAAACACAGCTCAAGACATCAACCCACTGTGCAGGCTTCATGGCCCAAG	5620
1761	ValGluGlnPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeuAla	1780
5621	GTAGAACAATTTCTGGGCCAAACACATGTGGAACTTCATTAGCGGCATCCAATACCTCGCA	5680
1781	GlyLeuSerThrLeuProGlyAsnProAlaValAlaIaSerMetMetAlaPheSerAlaAla	1800
5681	GGACTATCAACACTGCCAGGGACCTGCAGTAGCTTCATGATGGCGTTTCAGTGGCGGCC	5740
1801	LeuThrSerProLeuSerThrSerThrThrIleLeuLeuAsnIleLeuGlyGlyTrpLeu	1820
5741	CTCACCACTCCGCTGTCAACAGCACCACTATCTCTCAACATTTTGGGGGCTGGGCTA	5800
1821	AlaSerGlnIleAlaProProAlaGlyAlaThrGlyPheValValSerGlyLeuValGly	1840
5801	GCATCCCAAAATGCACCAACCGCGGGGCCACTGGCTTCGTGTTCAGTGGCTAGTGGGA	5860
1841	AlaAlaValGlySerIleGlyLeuGlyLysValLeuValAspIleLeuAlaGlyTyrGly	1860
5861	GCTCGCTAGCAGTATAGGCTTAGTAAAGTGCTAGTAGGACATCTCTGGCAGGGTATGGT	5920
1861	AlaGlyIleSerGlyAlaLeuValAlaPheLysIleMetSerGlyGluLysProSerMet	1880
5921	GGGGGCATTTGGGGGCTCTCGTCGCATTCAGGATCATGTCTGGCGAGAAGCCCTCCATG	5980
1881	GluAspValValAsnLeuLeuProGlyIleLeuSerProGlyAlaLeuValValGlyVal	1900
5981	GAGGATGCTCAACTTCTGCTCGCTGGAAATTCGTCTCCGGGTGCCCTTGGTAGTGGAGTC	6040
1901	IleCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrpMet	1920
6041	ATCTCGCGCGCATTTCTCGCGCGACACAGTGGGACCGGGGGAGGCGCGCTCCAAATGGATG	6100
1921	AsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaProThrHisTyrValThr	1940
6101	AATAGACTATTGGCTTTGCTTCCAGAGGAATACAGTCGCGCCCCCACCACCTACGTGAGC	6160
1941	GluSerAspAlaSerGlnArgValThrGlnLeuLeuGlySerLeuThrIleThrSerLeu	1960
6161	GAGTCGGATGCTCGCAGCGGTGACCAACTACTTTGGCTCCCTTTACCATACCGACCTG	6220
1961	LeuArgArgLeuHisAsnTrpIleThrGluAspCysProIleProCysGlySerTrp	1980
6221	CTCAGAAAGACTCCAACATGGGATTAATCAGGACTCGCCCATCCCATGGCGCGCTCGTGG	6280
1981	LeuArgAspValTrpAspTrpValCysThrIleLeuThrAspPheLysAsnTrpLeuThr	2000
6281	CTCCCGGATGTGTGGGACTGGGGTTTGACCATCTCTAACAGACTTTAAAAATTTGGCTGACC	6340
2001	SerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTyrLys	2020
6341	TCCAAATATTCCCAAGATGCCGCGCTCCCTTTGTCTCTGTCAAAAGGGGTACAAG	6400

Qy	2021	GlyValTrpAlaGlyThrGlyIleMetThrThrArgCysProCysGlyAlaAsnIleSer	2040	Db	7481	TTGGCCCTTTTCGAGACAGGTTCCATCTCTTCATGCCCCCTTCGAGGGGAGCTTGG	7540
Db	6401	GGCGTGGGGCGGCACTGGCATCATGACACACGCGTGTCTTGGCGGCGCAATATCTCT	6460	Qy	2401	AspProAspLeuGluProGluGlnValGluProGlnProGlnProGlnGlyGlyValAla	2420
Qy	2041	GlyAsnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTrp	2060	Db	7541	GATCCAGACTGAGGCTGAGCAGGTAGAGCCCAACCCCCCGGAGGGGGTGGCA	7600
Db	6461	GGCAATGTCCGTGGGCTCATGAGAAATCACGGGGCTTAAGACCTGCGTGAATATCTGG	6520	Qy	2421	AlaProGlySerAspSerGlySerTrpSerThrCysSerGluGluAspSerValVal	2440
Qy	2061	GlnGlyThrPheProIleAsnCysThrGluGlyGlnCysValProLysProAlaPro	2080	Db	7601	GCTCCCGGCTCGGACTCGGGTCTTGTCTTACTTGTCTCGAGGAGGACGACTCCGTCGTG	7660
Db	6521	CAGGGACCTTCTCATCAATTTGTACACGGAGGCGCAGTGGCGCAACCGCGCCA	6580	Qy	2441	CysCysSerMetSerTrpThrGlyAlaLeuIleThrProCysSerProGluGlu	2460
Qy	2081	AspPheLysValAlaIleTrpArgValAlaAlaSerGluTrpAlaGluValThrGlnHis	2100	Db	7661	TGCTGTCTCATGTCTACTCTCGACCGGGCTCTAATAACTCTCTTGTAGTCCGAGAG	7720
Db	6581	AACTTTAAGGTCCCATCTCGAGGCTGGCGGCTCAGAGTACCGGAGGTGACGACGAC	6640	Qy	2461	GluLysLeuProIleAsnProLeuSerAsnSerLeuLeuArgTyrrHisAsnLysValTyr	2480
Qy	2101	GlySerTyrrHisTyrrIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeu	2120	Db	7721	GAGAAGTTACCATTAACCCCTTGAGCAACTCTCTTGTGGATATCACAAAGGTGTAC	7780
Db	6641	GGGTATACCCACTACATACAGGACTCACCACTGATTAAGTTCCTCCCTGCCAACA	6700	Qy	2481	CysThrThrThrLysSerAlaSerLeuArgAlaLysLysValThrPheAspArgMetGln	2500
Qy	2121	ProSerProGluPhePheSerTrpValAspGlyValGlnIleHisArgPheAlaProThr	2140	Db	7781	TGTACCACAAACAAAGAGCGCTCACTAAGGGCTAAAGAGTAACCTTTTGTAGATGCAA	7840
Db	6701	CCCTCTCCCGAGTCTTTCTCTGGTGGACGGAGTGCAGATCCATAGTTTGGCCCCACA	6760	Qy	2501	ValLeuAspSerTyrrTyrrAspSerValLeuLysAspIleLysLeuAlaLysLysVal	2520
Qy	2141	ProLysProPhePheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal	2160	Db	7841	GTGCTCGACTCTCTACTACGACTCAGTCTTAAAGGACATTAAGCTAGCGGCTCCAAGTC	7900
Db	6761	CCGAAGCCGTTTTTCCGGGATGAGTCTCTGTTCTGCGTTCGGCTTAATTCATTTGTCGC	6820	Qy	2521	ThrAlaArgLeuLeuThrMetGluAlaCysGlnLeuThrProProHisSerAlaArg	2540
Qy	2161	GlySerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThr	2180	Db	7901	ACCGCAAGGCTCTCTACCATGAGGAGGCTTGCAGTTAACCCCACTTCTGCAAGA	7960
Db	6821	GGGTCCCGAGTCTCTTCCGACCCCTGAAACCCGACACAGACATTTGATGTCCATGCTAAC	6880	Qy	2541	SerLysTyrrGlyPheGlyAlaLysGluValArgSerLeuSerGlyArgAlaValAsnHis	2560
Qy	2181	AspProSerHisIleThrAlaGluThrAlaAlaArgArgLeuAlaArgGlySerProPro	2200	Db	7961	TCTAATATGGGTTTGGGCTAAGGAGTCCGAGCTTCTCGGAGGCGGCTTAACCA	8020
Db	6881	GATCCATCTCATATCACGGCGGAGACTGACGCGGGGCTTTAGCGGGGGTCAACCCCA	6940	Qy	2561	IleLysSerValTrpLysAspLeuLeuGluAspSerGluThrProIleProThrThrIle	2580
Qy	2201	SerGluAlaSerSerAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThr	2220	Db	8021	ATCAAGTCCGTGGAAGGACCTCTCTGGAGGACTCAGAAACCACTTCCCAACCAAT	8080
Db	6941	TCCGAGCAAGCTCTCGCGGAGCCAGCTATCGGACCATCGCTGCGAGCCACTGCACC	7000	Qy	2581	MetAlaLysAsnGluValPheCysValAspProThrLysGlyGlyLysValAlaArg	2600
Qy	2221	ThrHisGlyLysAlaTrpAspValAspMetValAspAlaAsnLeuPheMetGlyGlyAsp	2240	Db	8081	ATGGCCAAAATATGAGTGTCTTGGCGTGGAGCCCAAGGGGGGCAAGAAAGCAGCTCG	8140
Db	7001	ACCCACGGCAAGGCTATGATGTGGACATGGTGGATGTCTAACCCTGTTTCATGGGGGCGAT	7060	Qy	2601	LeuIleValTyrrProAspLeuGlyValArgValCysGluLysMetAlaLeuTyrrAspIle	2620
Qy	2241	ValThrArgIleGluSerGlySerLysValValValLeuAspSerLeuAspProMetVal	2260	Db	8141	CTTATCGTTTACCTTGACCTCGGCTCGGCTGCGAGAGATGCGCCCTTTATGACAT	8200
Db	7061	GTGACTCGGATAGAGTCTGGGTCAAGTGGTCTCTGAGCTCTCTCGACCCCAATGGTC	7120	Qy	2621	ThrGlnLysLeuProGlnAlaValMetGlyAlaSerTyrrGlyPheGlnTyrrSerProAla	2640
Qy	2261	GluGluArgSerAspLeuGluProSerIleProSerGluTyrrMetLeuProLysLysArg	2280	Db	8201	ACACAAAATCTCTCAGCGGCTGATGGGGCTTCTTATGGATTCCAGTATTCGCCCGCT	8260
Db	7121	GAAGAAAGGAGGACCTTGAGCCTTCGATACCATCAGATACATGCTCCCAAGAGAGG	7180	Qy	2641	GlnArgValGluPheLeuLeuLysAlaTrpAlaGluLysLysAspProMetGlyPheSer	2660
Qy	2281	PheProProAlaLeuProAlaTrpAlaArgProAspTyrrAsnProProLeuValGluSer	2300	Db	8261	CAGCGGTAGAGTTCTCTTGAAGCATGGGGGCAAGAGACCTTATGGGTTTTTCG	8320
Db	7181	TTCCCAACAGCTTTTACCGGCTCGGCAAGGCTGATTACACCCACCGCTTGTGGAATCG	7240	Qy	2661	TyrAspThrArgCysPheAspSerThrValThrGluArgAspIleArgThrGluGluSer	2680
Qy	2301	TrpLysArgProAspTyrrGlnProAlaThrValAlaGlyCysAlaLeuProProProArg	2320	Db	8321	TATGATACCGATGCTTTGACTCAACCTGCTGAGAGAGACATCAGGACTGAGGAGTCC	8380
Db	7241	TGGAAAGGCGAGATACCAACCGCCACTTGTTCGGGGCTGTCTCTCTCTCTCTCTAGG	7300	Qy	2681	IleTyrrArgAlaCysSerLeuProGluAlaHisThrAlaIleHisSerLeuThrGlu	2700
Qy	2321	LysThrProThrProProProArgArgArgThrValGlyLeuSerGluAspSerIle	2340	Db	8381	ATATATCGGGCTGCTCTCTTGCAGAGGAGGCCACATGCCATACATCTCGCTAACTGAG	8440
Db	7301	AAAACCCGAGCGCTCCCGCAAGGAGGCGCGGACAGTGGGCTTAAGTAGGACTCCATA	7360	Qy	2701	ArgLeuTyrrValGlyProMetPheAsnSerLysGlyGlnThrCysGlyTyrrArgArg	2720
Qy	2341	GlyAspAlaLeuGlnLeuAlaIleLysSerPheGlyGlnProProProSerGlyAsp	2360	Db	8441	AGACTTTACGTGGAGGGCTTATGTTCAACAGCAAGGGGCAACCTGCGGGTACAGCGCT	8500
Db	7361	GGAGATGCCCTTCAACAGCTGGCCATTAAGTCTCTTGGCCAGGCGCCCCCAAGCGCGAT	7420	Qy	2721	CysArgAlaSerGlyValLeuThrThrSerMetGlyAsnThrIleThrCysTyrrValLys	2740
Qy	2361	SerGlyLeuSerThrGlyAlaGlyAlaAlaAspSerGlySerGlnThrProProAspGlu	2380	Db	8501	TGCCCGCCAGCGGGGTGCTCACCATGACTAGCATGGGGAACCATCATCATCTACGTGAAA	8560
Db	7421	TCAGGCTCTTCCAGGGGGGGCGCTGCCGATTCGGGAGTCCAGAGCGCTCTCTGTAGAG	7480	Qy	2741	AlaLeuAlaAlaCysLysAlaAlaGlyIleIleAlaProThrMetLeuValCysGlyAsp	2760
Qy	2381	LeuAlaLeuSerGluThrGlySerIleSerMetProProLeuGluGlyLeuGly	2400				

8561 GCCTTAGCGGCTTTAAAGCTGAGGATAATCGCGCCCAACATGCTGGTATCGCGCAT 8620
 2761 AspLeuValVallleSerGluSerGlnGlyThrGluGluAspGluArgAsnLeuArgAla 2780
 8621 GACTTGGTTCTCATCTCAGAAAGCCAGGAGACCGAGGAGACCGAGCGAACCTGAGAGCC 8680
 2781 PheThrGluAlaMetThrArgTySerAlaProProGlyAspProProArgProGluTy 2800
 8681 TTCAGGAGGCTATGACAGGATTTCTGCCCTCTCTGCTGACCCCGCCAGCCGAGAT 8740
 2801 AspLeuGluLeuLeuThrSerCysSerAsnValSerValAlaLeuGlyProGlnGly 2820
 8741 GATCTGAGCTGATAACATCTTCTCTCAATGTGTCTGTGGCTGGGCCCCACAGGC 8800
 2821 ArgArgGlyTyThrLeuThrArgAspProThrProLeuAlaArgAlaAlaTrpGlu 2840
 8801 CGCGCAGATACCTGACGAGACCTTACCATCTCCATTCCTGCTGCTGCTGCTGCTG 8860
 2841 ThrValArgHisSerProValAsnSerTrpLeuGlyAsnLeuLeuGlnTyAlaProThr 2860
 8861 ACAGTTAGACATCTCCCTGTCAATTCATGCTGGGAAACATCATCCAGTACGCCCGACC 8920
 2861 IleTrpAlaArgMetValLeuMetThrHisPhePheSerLeuMetAlaGlnAspThr 2880
 8921 ATATGGGCTCGCATGGTCTGATGACACATCTTCTTCTCCATTCATGGCTCAAGACAG 8980
 2881 LeuAspGlnAsnLeuAsnPheGluMetTyGlyAlaValTySerValSerProLeuAsp 2900
 8981 CTGGACAGAACCTCAACTTTGAGATGACGAGCGGTGACTCCGCTGCTGCTGCTGCTG 9040
 2901 LeuProAlaIleLeuGluArgLeuHisGlyLeuAspAlaPheSerLeuHisThrTyThr 2920
 9041 CTCCAGCTATATTAAGGTTACATGGCTTGACGCTTTCTTCTGACACATACACT 9100
 2921 ProHisGluLeuThrArgValAlaSerAlaLeuArgLysLeuGlyAlaProProLeuArg 2940
 9101 CCCACAGAACTGACACGGGTGGCTTACGCTCAGAAACCTTGGGGCGCCACCCCTCAGA 9160
 2941 AlaTrpLysSerArgAlaArgAlaValArgAlaSerLeuLeuSerArgGlyGlyArgAla 2960
 9161 GCGTGAAGAGCGCGGACGTCAGTACGAGCGCTCCCTCATCTCCGCTGGGGGAGAGCG 9220
 2961 AlaValCysGlyArgTyLeuPheAsnTrpAlaValLysThrLysLeuLysLeuThrPro 2980
 9221 GCCGTTTGGGTCGATATCTTCAATTTGGCGGTGAAGCAAGCTCAAACTCACTCCA 9280
 2981 LeuProGluAlaArgLeuLeuAspLysSerTrpPheThrValGlyAlaGlyGlyGly 3000
 9281 TTGCGGAAAGCGCGCTCTCTGATTTATCCAGCTGGTTACCGTGGCGCGCGCGGGC 9340
 3001 AspileTyHisSerValSerArgAlaArgProArgLeuLeuPheGlyLeuLeuLeu 3020
 9341 GACATTTATCAACGCTGTCGCGTGGCGGCGGACCCCGCTTATTTGCTTTTGGCTACTCCTA 9400
 3021 LeuPheValGlyValGlyLeuPheLeuLeuProAlaArg 3033
 9401 CTTTTTGTAGGGTAGGCGCTTCTTCTACTCCCGCTCGG 9439

RESULT 2
 AX057317
 LOCUS 9711 bp DNA linear PAT 17-JAN-2001
 DEFINITION Sequence 3 from Patent WO0075352.
 AX057317
 VERSION AX057317.1 GI:12310062
 KEYWORDS
 SOURCE Hepatitis C virus
 ORGANISM Hepatitis C virus
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 Hepacivirus.
 1
 REFERENCE
 AUTHORS Nam, J.H., Bukh, J., Emerson, S.U. and Purcell, R.H.
 TITLE Hcv/bv/dv chimeric genomes and uses thereof
 JOURNAL Patent: WO 0075352-A 3 14-DEC-2000;

THE GOVERNMENT OF THE UNITED STATES OF AMERICA (US)
 Location/Qualifiers
 1..9711
 /organism="Hepatitis C virus"
 /mol_type="genomic DNA"
 /db_xref="taxon:11103"
 BASE COUNT 1987 a 2853 c 2668 g 2203 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 9711
 Score: 16120.00 Matches: 3033
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-980-559-2 (1-3033) x AX057317 (1-9711)
 QY 1 MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGln 20
 Db 341 ATGAGCACAATCTTAACCTCAAGAAAACCAAAAGAAAACACCAACCGTCCGCCACAA 400
 QY 21 AspValLysPheProGlyGlyGlnIleValGlyGlyValTyrlleuLeuProArgArg 40
 Db 401 GACGTTAAGTTTCGGGCGCGCCAGATCGTTGGCGAGTATACTTGTTCGCCGCGAGG 460
 QY 41 GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly 60
 Db 461 GCGCCAGGTGGGTGTGCGCGCAGCAAGAAAGACTTCGGAGCGGTCCCGCCAGCGTGA 520
 QY 61 ArgArgGlnProLysProLysAspArgArgSerThrGlyLysSerTrpGlyLysProGly 80
 Db 521 AGCGCCAGCCATCCCTAAGATCGCGCTCCACTGGCAATCTCTGGGAAACACAGGA 580
 QY 81 TyrProTrpProLeuTyrlleGlyAsnGluGlyLeuGlyTyrlleAlaGlyTrpLeuLeuSerPro 100
 Db 581 TACCCCTGGCCCTATACCGGAATGAGGACTCGGCTGGCGAGGATGGCTCTCTGCCCC 640
 QY 101 ArgGlySerArgProSerTrpGlyProAsnAspProArgHisArgSerArgAsnValGly 120
 Db 641 CGAGGTTCCCGTCCCTCTTTGGGCGCCCAATGACCCCGGCATAGGTCGGCGCAACGTGG 700
 QY 121 LysValLysAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrlleProValVal 140
 Db 701 AAGTCAATCATACCTTAACGTCGGCTTTGCCGACTCATGGGGTACATCTCTCTGCTG 760
 QY 141 GlyAlaProLeuGlyGlyValAlaArgAlaLeuAlaHisGlyValArgValLeuGluAsp 160
 Db 761 GCGCGCCCGCTCGCGCGCTCGCCAGAGCTCTCGCGCATGGCGTGGAGAGTCTCTGGAGG 820
 QY 161 GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerilePheLeuLeuAla 180
 Db 821 GGGGTTAAATTTGCAACAGGGAACATTACCGGTTGCTCTCTTTCTATCTTTCTGTGGCC 880
 QY 181 LeuLeuSerCysileThrThrProValSerAlaAlaGluValLysAsnIleSerThrGly 200
 Db 881 CTGCTGCTGCTCATCACCCCGCTCTCGCTGCGCGAGTGAAGAACATCACTACGCGGC 940
 QY 201 TyrMetValThrAsnAspCysThrAsnAspSerileThrTrpGlnLeuGlnAlaVal 220
 Db 941 TACATGGTACTAACGACTGACCAATGACAGCATTTACCTGGCAGCTCCAGGCTGCTGTC 1000
 QY 221 LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTrpIle 240
 Db 1001 CTCACGTCGCCGGGTGCGTCCCGTCGAGAAAGTGGGGAATGCATCTCAGTCTCGATA 1060
 QY 241 ProValSerProAsnValAlaValGlnArgProGlyValAlaLeuThrGlnGlyLeuArgThr 260
 Db 1061 CCGGTCTCACCGAATGTGGCGGTGCGCGGCCCGGCCCTCACGAGGCGCTTGGGACG 1120
 QY 261 HisileAspMetValValMetSerAlaThrLeuLeuCysSerAlaLeuTyrlleValGlyAspLeu 280

1121 CACATCGACATGGTGTGATGTCGCCACCGCTCTGCTCTCCCTCTACGTGGGGACCTC 1180
 281 CysGlyGlyValMetLeuAlaAlaGlnMetPheIleValSerProGlnHisHisLeuPhe 300
 1181 TGGGTGGGTGATGCTCGAGCCCAATGTTCAATGTTCTCGCGCAGCACCACTGGTTT 1240
 301 ValGlnAspCysAsnCysSerIleTyrProGlyThrIleThrGlyHisArgMetAlaTyr 320
 1241 GTCCAAAGACTGCAATTCCTCATCTACCTGGTACCATCATCTGGACACCGCATGGCATG 1300
 321 AspMetMetMetAsnTyrProThrAlaThrMetIleLeuAlaTyrAlaMetArgVal 340
 1301 GACATGATGATGAACCTGGTCCGCCACCGCTACCATGATCTTGGCGTACGGATGCGGTG 1360
 341 ProGluValIleIleAspIleIleSerGlyAlaHisTyrGlyValMetPheGlyLeuAla 360
 1361 CCCGAGGTCAATATAGACATCATTAGCGGGCTCATTTGGGGCTCATGTTGGCTTGGCC 1420
 361 TyrPheSerMetGlnGlyAlaTyrAlaLysValValIleLeuLeuAlaAlaGly 380
 1421 TACTTCTCTATGCGAGGAGCGTGGCGAAGTCTGTCTCATCTCTGTTGGCGCGCGG 1480
 381 ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr 400
 1481 GTGGACCGCGCACCATCTACTGTTGGGGGTCTGCGCGCAGACCAACCGGGCGCTCAC 1540
 401 SerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTyr 420
 1541 AGCTTATTTGACATGGGCCCCAGGAGAAATCCAGCTCGTTAACCAATGGCAGCTGG 1600
 421 HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer 440
 1601 CACATCAACCGCACCGCCCTGAACTGCAATGACTCTCTTGACACCGGCTTTATCGGCTC 1660
 441 LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg 460
 1661 CTGTCTTACACCCACACAGCTTCAACTCGTCAGAGTGTCCCGAACGCATGTCCGCTGCC 1720
 461 SerIleGluAlaPheArgValGlyTyrGlyAlaLeuGlnTyrGluAspAsnValThrAsn 480
 1721 AGTATCGAGGCTTCCGGGTGGAGTGGGGCGCTTGCATATAGAGTATATGACCAAT 1780
 481 ProGluAspMetArgProTyrCysTyrPheIleTyrProProArgGlnCysGlyValValSer 500
 1781 CCAGAGCATATGAGACCTTATGTCGACTACCCACCAAGGCGAGTGGCGGTCTCC 1840
 501 AlaIysThrValCysGlyProValTyrCysPheThrProSerProValValValGlyThr 520
 1841 GCGAAGACTGTGTGGCCCCAGTGACTGTCTTCAACCCCGACCGCAGTGTGTGGGCACG 1900
 521 ThrAspArgLeuGlyAlaProThrTyrThrTyrGlyGluAsnGluThrAspValPheLeu 540
 1901 ACCGACAGGCTTGGAGCGCCACTTACAGTGGGGGAGAAATGAGACAGATGTTCTCCCTA 1960
 541 LeuAsnSerThrArgProProLeuGlySerTyrPheGlyCysThrTyrMetAsnSerSer 560
 1961 TTGAACAGCACTCGACACCGCTGGGTCTGCTGCTGGCTGACCTGGATGAATCTTCT 2020
 561 GlyTyrThrIysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer 580
 2021 GGCTACACCAAGACTTGGCGCGCACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2080
 581 ThrAspLeuLeuCysProThrAspCysPheArgIleHisProAspThrThrTyrLeuLys 600
 2081 ACGGACCTGTGTGGCCCCACGGACTGTTTAGAAGCATCTCTATACCATCTTACCTCAA 2140
 601 CysGlySerGlyProTyrLeuThrProArgCysLeuIleAspTyrProTyrArgLeuTyr 620
 2141 TGGCGCTCTGGGCGCTGCTCAGCCCAAGGTGCTGATCATCTACCTACAGGCTCTGG 2200
 621 HisTyrProCysThrValAsnTyrThrIlePheLysIleArgMetTyrValGlyVal 640
 2201 CATTAACCTTGCACAGTTAACTATACCATCTTCAAAATAAGGATGTATGTGGGAGGGTT 2260

641 GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp 660
 2261 GAGCACAGGCTCAGCGCTGATGCAATTTCACTCGTGGGATCGTTGCAACTTGGAGGAC 2320
 661 ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGluTyrPalalleLeuPro 680
 2321 AGAGACAGAAGTCAACTGTCTCTCTTTGTTGTCATCCACACCGAATGGGCCATTTTACCT 2380
 681 CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisLeuHisGlnAsnIle 700
 2381 TGTCTTACTCGACCTGCCCTTGTGACTGGTCTCTCCACCTCCACCAAAACATC 2440
 701 ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrIlysTyrIleValArgTyr 720
 2441 GTGACGTGACAAATTCATGATATGCGCTATCACCTGCCCTCACAAATATACGTCGATG 2500
 721 GluTyrValIleLeuLeuPheLeuLeuAlaAspAlaArgValCysAlaCysLeuTyr 740
 2501 GAGTGGGTAAATCTCTTATTTCTGCTCTTAGCGGACGCCAGGCTTGGCGCTGCTTATGG 2560
 741 MetLeuIleLeuLeuGlyGlnAlaGluAlaAlaLeuGluLysLeuValIleLeuHisAla 760
 2561 ATGCTCATCTGTGTGGGCCAGGCCGAGCAGCAGCAGCTAGAGAGCTGTGCTATCTTGCAC 2620
 761 AlaSerAlaAlaSerCysAsnGlyPheLeuTyrPheValIlePheValAlaAlaTyr 780
 2621 GCGAGCGAGCTAGCTGCAATGGCTTCTATATTTTGTCTATCTTTTGTGGCTGCTTGG 2680
 781 TyrIleLysGlyArgValValProLeuAlaThrTyrSerLeuThrGlyLeuTyrSerPhe 800
 2681 TACATCAAGGTCGGGTAGTCTCCCTTAGCTACCTATTCCTCCTACCTGCTGCTGCTCTT 2740
 801 SerLeuLeuLeuAlaLeuProGlnGlnAlaTyrAlaTyrAspAlaSerValHisGly 820
 2741 AGCTTACTCTCTAGCATGTGCCCAACAGGCTTATGCTTATGACCATCTGTGCAATGGC 2800
 821 GlnIleGlyAlaAlaLeuLeuValMetIleThrLeuPheThrLeuThrProGlyTyrLys 840
 2801 CAGATAGAGAGCGCTCTGCTGGTATGATCATCTCTTACTCTACCCCGGCTATAG 2860
 841 ThrLeuLeuSerArgPheLeuTyrTrpLeuCysTyrLeuLeuThrLeuGlyGluAlaMet 860
 2861 ACCCTTCTCAGCGGTTTGTGGTGTGCTATCTTCTGACCTGCGGGAAGCTATG 2920
 861 ValGlnGluTyrAlaProProMetGlnValArgGlyGlyArgAspGlyIleIleTyrAla 880
 2921 GTCCAGGAGTGGGCACCATCTATGACAGGTGCGCGGTGCGTATGATGCGCATCATATG 2980
 881 ValAlaIlePheTyrProGlyValValPheAspIleThrLysTyrLeuLeuAlaValLeu 900
 2981 GTCCCATATTTTACCCAGGTGTGGTGTGATGATGATGATGATGATGATGATGATGATG 3040
 901 GlyProAlaTyrLeuLeuLysGlyAlaLeuThrArgValProTyrPheValArgAlaHis 920
 3041 GGGCTGTCTTACTCTCTAAAGGTGCTTTGACGCGGTGCGGTACTTCTGTCAGGCTCAC 3100
 921 AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyGlyArgTyrValGlnMet 940
 3101 GCTTACTAGGATGTGCACCATGTCAGAGGCATCTCGCGGGGGGAGGACGTGCTCCAGATG 3160
 941 AlaLeuLeuAlaLeuGlyArgTyrThrGlyThrTyrIleTyrAspHisLeuThrProMet 960
 3161 GGTCTACTAGCCCTTGGCAGGTGGACTGGCATCTATCATGACCATCTACCCCTATG 3220
 961 SerAspTyrAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIleIlePhe 980
 3221 TCGGATGGGCTGTAGTGGCTGCGGACCTGCGGCTGCGGCTGCGGCTGAGCTATCATCTTC 3280
 981 SerProMetGluLysLysValIleValTyrGlyAlaGluThrAlaAlaCysGlyAspIle 1000
 3281 AGTCCGATGAGAAAGTCAATGTCTGGGGAGGCGAGACAGCTGCTTGTGGGGACATT 3340

1001 LeuHisGlyLeuProValSerAlaArgLeuGluValLeuLeuGlyProAlaAsp 1020
 3341 TTACAGGAGCTTCCCGTGTCCGCGGACTTGGTGGGAGGTCTCTCTGGCCAGCTGAT 3400
 1021 GlyTyrThrSerLysGlyTyrSerLeuLeuAlaProIleThrAlaTyrAlaGlnThr 1040
 3401 GGCTATACCTCCAAAGGGGTGGAGTCTCTCGCCCGCATCACTGCTTACGCCACGACACA 3460
 1041 ArgGlyLeuLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGluGlnAla 1060
 3461 CGTGGCCCTTTGGGACCATAGTGGTGGAGCATGACGGGGCGGACAGAACAGAGCT 3520
 1061 GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly 1080
 3521 GGGGAATTCAGGTCCTGTCCACAGTCACCTCAGTCTCTCTCGGAACATCCATCTCGGG 3580
 1081 ValLeuTrpThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro 1100
 3581 GTTTGTGGACTGTCTACCATGGAGCTGGCAACAGACTCTGGCCGGCTCACGGGGTCCG 3640
 1101 ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTrpProSerProGly 1120
 3641 GTACGCGAGATGTACTCAGTGTGGAGGGGACTTAGTAGGGTGGCCAGCCCCCTGGG 3700
 1121 ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArgAsn 1140
 3701 ACTAATCTTTGGAGCCGTGACAGTGTGGAGCGGTGACCTGTACCTGTGTACGCGGAAC 3760
 1141 AlaAspValIleProAlaArgArgGlyAspLysArgGlyAlaLeuLeuSerProArg 1160
 3761 GCTCATGTCTCCCGGCTCGAAGACGCGGGGACAAACGGGGAGCGCTACTCTCCCGAGA 3820
 1161 ProLeuSerThrLeuLysGlySerSerGlyGlyProValLeuLysProArgGlyHisAla 1180
 3821 CCTCTTTCACCTTGAAGGGGTCTCAGAGGGCCCGGTGTATGCCCGGCGGCGACGCT 3880
 1181 ValGlyValPheArgAlaValCysSerArgGlyValAlaLysSerIleAspPheIle 1200
 3881 GTGGAGTCTTCCGGGAGCTGTGTCTCTCGGGCGTGGCTAAGTCCATAGATTTCATC 3940
 1201 ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro 1220
 3941 CCCGTTGAGACACTCGACATCGTCACCGGTCGCCACCTTTAGTGAACAACAGCACCA 4000
 1221 ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLys 1240
 4001 CCTGCTGTGCCCGGAGCTATCAGGTGGGTACTTGTATGCCCGGAGCTGGAGTGGAAAG 4060
 1241 SerThrLysValProValAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnPro 1260
 4061 AGCACCAAGTCTCTGTCCGATATGCTGTGCTCAGGGGTATAAAGTGTAGTGTATATCCC 4120
 1261 SerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerLysAlaHisGlyIleAsnPro 1280
 4121 TCAGTGGCTGCCACCTCGGGGTTTGGGCGGTACTTGTCTAAGGCACATGGCATCAATCCC 4180
 1281 AsnIleArgThrGlyValArgThrValThrThrGlyAlaProIleThrTyrSerThrTyr 1300
 4181 AACATTAGGACTGGAGTCAGACTGTGACAGCCGGGGCGGCCATCAGTACTCCACATAT 4240
 1301 GlyLysPheLeuAlaAspGlyGlyCysAlaGlyValTyrAspIleIleCysAsp 1320
 4241 GGCAATTCCTCGCGGATGGGGGTGTGGGGCGGCGGCTTACGACATCATCATATGTGAT 4300
 1321 GluCysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGlnAla 1340
 4301 GAATGCCATGCGGTGGACTCTACCATCATCTTGGCATCGGAACAGTCTTGTATCAAGCA 4360
 1341 GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProProGlySerValThr 1360
 4361 GAGACAGCTGGGGTCAGACTAATCTGTGCTGCTACAGCTACGCCCCCTCGGTTCAGTGACA 4420
 1361 ThrProHisProAsnIleGluGluValAlaLeuGlyGlnGluGlyGluIleProPheTyr 1380

4421 ACCCCCAACCCCAACATAGAGAGGTGGCCCTTGGGCGAGGGCGAGATCCCTCTCTAT 4480
 1381 GlyArgAlaIleProLeuSerTyrIleLysGlyArgGlyHisLeuIlePheCysHisSer 1400
 4481 GGGAGGGCGATTCCCTGTCTTACATCAAGGGAGGAAGACATCTGATCTTCTGCCATTCA 4540
 1401 LysLysLysCysAspGluLeuAlaAlaLeuArgGlyMetGlyLeuAsnSerValAla 1420
 4541 AAAAAAAGTGTGACGAGCTCGGGGGCCCTTGGGGTATGGCTTGAACCTCAGTGGCA 4600
 1421 TyrTyrArgGlyLeuAspValSerValIleProThrGlnGlyAspValValValAla 1440
 4601 TACTACAGAGGGTGGAGCTCTCCGTAATACCAACTCAGGAGAGCTAGTGTGTGGTGGC 4660
 1441 ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnVal 1460
 4661 ACCGACGCGCTCATGACAGGGTATATCGGGGACTTTGATCTCGGTATCGACTGCAACGTA 4720
 1461 AlaValThrGlnValValAspPheSerLeuAspProThrPheThrIleThrGlnIle 1480
 4721 GGCTCACTCAAGTTGTAGACTTCAGTTTACACCCACACATTCCATATACACACAGATT 4780
 1481 ValProGlnAspAlaValSerArgSerGlnArgGlyArgThrGlyArgGlyArgLeu 1500
 4781 GTCCCTCAAGACGCTGTCTCAGTAGCAGCGCGGGTCCGACCGGTAGGGGAGACTG 4840
 1501 GlyIleTyrArgTyrValSerThrGlyGluArgAlaSerGlyMetPheAspSerValVal 1520
 4841 GGCATTATAGGTATGTTCCTCAGTGGAGCGGCTCAGGAATGTTTACAGTAGTGTAGTG 4900
 1521 LeuCysGluCysTyrAspAlaGlyAlaAlaTyrTyrGluLeuThrProSerGluThrThr 1540
 4901 CTCTGTAGTGTACGACGAGGGGCGCATGTGTATGAGCTCACACCATCGGAGACCACC 4960
 1541 ValArgLeuArgAlaTyrPheAsnThrProGlyLeuProValCysGlnAspHisLeuGlu 1560
 4961 GTCAGGCTCAGGGCGTATTTCAACACGCGCGTGTGCTGTGTGCCAAGACCATCTTGAG 5020
 1561 PheTrpGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThr 1580
 5021 TTTTGGGAGGAGTGTTCACCGCGCTCACACATAGATGCCCATCTCTCTTCCCAACA 5080
 1581 LysGlnSerGlyGluAsnPheAlaTyrLeuThrAlaTyrGlnAlaThrValCysAlaArg 1600
 5081 AAGCAATCGGGGAAATTTTCGATATTAACAGCTTACAGGCTACAGGTGTGGCTAGG 5140
 1601 AlaLysAlaProProProSerTrpAspValMetTrpLysCysLeuThrArgLeuLysPro 1620
 5141 GCCAAAGCCCCCCCCCTCTGGGAGCTCATGTGGAAGTGTGTGACTCGACTCAAGCCCC 5200
 1621 ThrLeuValGlyProThrProLeuLeuTyrArgLeuGlySerValThrAsnGluValThr 1640
 5201 ACATCTGGGGCCCCACCTCTCTCCCGCTTGGGCTGTGTATACCAACAGGCTCACC 5260
 1641 LeuThrHisProValThrLysTyrIleAlaThrCysMetGlnAlaAspLeuGluValMet 1660
 5261 CTCACATCCCCGTGAGAAATACATCGCCACCTGATGCAAGCCGACCTTGAAGTCAATG 5320
 1661 ThrSerThrTrpValLeuAlaGlyGlyValLeuAlaAlaValAlaAlaTyrCysLeuAla 1680
 5321 ACCAGCACATGGTCTTGGCAGGGGAGTCTTGGCGCGCTCGCCGCTATTGCCCTGGCG 5380
 1681 ThrGlyCysValCysIleIleGlyArgLeuHisIleAsnGlnArgAlaValAlaPro 1700
 5381 ACCGGGTGTGTTCATCATCGCCGCTTGCATTAACACAGAGCGGCTGTGGCGCG 5440
 1701 AspLysGluValLeuTyrGluAlaPheAspGluMetGluCysAlaSerArgAlaAla 1720
 5441 GACAGAGGTCTCTATGAGGGCTTTTGTAGATGGAGGATGTGCTCTAGGGCGGCT 5500
 1721 LeuIleGluGlyGlnArgIleAlaGluMetLeuLysSerLysIleGlnGlyLeuLeu 1740

Db	5501	CTCATTTGAAGAGGGGCGAGTAGCCGAGATGCTGAAGTCCAAAGTCCAAAGGCTTATTG	5560	QY	2101	GlySerTyrHisTyrIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeu	2120
QY	1741	GlnGlnAlaSerLysGlnAlaGlnAspIleGlnProThrValGlnAlaSerTrpProLys	1760	Db	6641	GGGTATACCACTACATAACAGGACTCACCACTGATAACTTGAAAGTCCCTCCGCAACTA	6700
Db	5561	CAGCAAGCTTCCAAACAAAGCTCAAGACATACACCCACCTGTGCGAGCTTCATGGCCCAAG	5620	QY	2121	ProSerProGluPhePheSerTrpValAspGlyValGlnIleHisArgPheAlaProThr	2140
QY	1761	ValGluGlnPheTrpAlaLysHisMetTrpAsnPhelIleSerGlyIleGlnTyrLeuAla	1780	Db	6701	CCCTCTCCCGAGTTCTTTTCTCGGTGGAGCGGAGTGCAGATCCATAGGTGTGGCCCA	6760
Db	5621	GTAGAACAAATTCCTGGGCGCAACACATGTGGAACTTCATTAGCGGCATCCAAATACCTCGCA	5680	QY	2141	ProLysProPhePheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal	2160
QY	1781	GlyLeuSerThrLeuProGlyAnnProAlaValAlaSerMetMetAlaPheSerAlaAla	1800	Db	6761	CCGAAGCCGTTTTCGGGATGAGGTCTCGTTCGTGGTGGGCTTAATTCATTGTGTCG	6820
Db	5681	GGACTATCAACACTGCCAGGAACCCCTGCGAGTACCTTCATGATGGCGTTCACTGGCCGC	5740	QY	2161	GlySerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThr	2180
QY	1801	LeuThrSerProLeuSerThrSerThrIleLeuLeuAsnIleLeuGlyGlyTrpLeu	1820	Db	6821	GGGTCCCACTTCTTCGACCTCGAACCCGACACAGAGCTATTGATGTCCATGCTTAACA	6880
Db	5741	CTCACCACTCCGCTGTCAACAGCACCACTATCTCTCAACATTTTGGGGGCTGGCTA	5800	QY	2181	AspProSerHisIleThrAlaGluThrAlaAlaArgArgLeuAlaArgGlySerProPro	2200
QY	1821	AlaSerGlnIleAlaProProAlaGlyAlaThrGlyPheValValSerGlyLeuValGly	1840	Db	6881	GATCCATCTCATATCACGGCGGAGACTGCAGCGCGCGTATTAGCGCGGGGTCAACCCCA	6940
Db	5801	GCATCCCAAAATTCACCCCGGGGGCCCACTGGCTTCGTTGTCAGTGGCTTAGTGGGA	5860	QY	2201	SerGluAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThr	2220
QY	1841	AlaAlaValGlySerIleGlyLeuGlyLysValLeuValAspIleLeuAlaGlyTyrGly	1860	Db	6941	TCCGAGGCAAGCTCTCTCGCGAGCCAGCTATCGGCACCATCGCTGCGAGCCACTGCACC	7000
Db	5861	GCTGCCGTAGGCAGTATAGGCTTAGGTAAAGTGTCTAGTGACATCTCTGGCAGGGTATGGT	5920	QY	2221	ThrHisGlyLysAlaTyrAspValAspMetValAspAlaAsnLeuPheMetGlyGlyAsp	2240
QY	1861	AlaGlyIleSerGlyAlaLeuValAlaPheLysIleMetSerGlyGlyLysProSerMet	1880	Db	7001	ACCCACGGCAAGCCCTATGATGTGGACATGGTGGATGCTAACTGTTCATGGGGGGGAT	7060
Db	5921	CGGGGCATTTTCGGGGGCTCTGTCGCATTCAGATCATGTCTGCGCAGAAAGCCCTCATG	5980	QY	2241	ValThrArgIleGluSerGlySerLysValValValLeuAspSerLeuAspProMetVal	2260
QY	1881	GluAspValValAsnLeuLeuProGlyIleLeuSerProGlyAlaLeuValValGlyVal	1900	Db	7061	GTGACTCGGATAGAGTCTGGGTCCAAAGTGGTCTGTTCTCGACTCTCTCGACCCCAATGGTC	7120
Db	5981	GAGGATGTCGTAACCTTGCTGCTCGGAATTCCTCTCCGGGTGCCCTTGGTAGTGGGAGTC	6040	QY	2261	GluGluAspSerAspLeuGluProSerIleProSerGlyTyrMetLeuProLysLysArg	2280
QY	1901	IleCysAlaAlaIleLeuArgArgHisValGlyProGlyGlyAlaValGlnTrpMet	1920	Db	7121	GAAGAAAGGAGGAGCTTGAGCTTCGATACCATCAGAATACATGCTCCCAAGAAAGAGG	7180
Db	6041	ATCTGGCGGCCCATTTCTGGCCGACACAGTGGGACCGGGGAAGCGCGCTCCAAATGGATG	6100	QY	2281	PheProProAlaLeuProAlaTyrAspProAspTyrAsnProProLeuValGluSer	2300
QY	1921	AsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaProThrHisTyrValThr	1940	Db	7181	TTCCCAACCACTTACCGGCTTGGGACCGGCTGATTACACCCACCGCTTGTGGAAATCG	7240
Db	6101	AATAGACTATTCGCTTGTTCAGAGGAATACAGTGGCCCCCAACCACTACGTGACG	6160	QY	2301	TrpLysArgProAspTyrGlnProAlaThrValAlaGlyCysAlaLeuProProArg	2320
QY	1941	GluSerAspAlaSerGlnArgValThrGlnLeuLeuGlySerLeuThrIleThrSerLeu	1960	Db	7241	TGGAAGGAGCCAGATTACCAACCGGCACTGTGGGGCTGTGCTCTCCCTCTCTCTAGG	7300
Db	6161	GAGTCGATGCGTCGACGCTGTGACCAACTACTTGGCTCCCTTACCAATACCAACGCTG	6220	QY	2321	LysThrProThrProProArgArgArgThrValGlyLeuSerGluAspSerIle	2340
QY	1961	LeuArgArgLeuHisAsnTrpIleThrGluAspCysProIleProCysGlyGlySerTrp	1980	Db	7301	AAAAACCCGACCCCTCCCAAGGAGGCGCGGACAGTGGGCTTAAGTGAGGACTCCATA	7360
Db	6221	CTCAGAAGACTCCACAACTGGATTACTGAGGACTGCCCATCCCATGCGCGGCTCGTGG	6280	QY	2341	GlyAspAlaLeuGlnGlnLeuAlaIleLysSerPheGlyGlnProProProSerGlyAsp	2360
QY	1981	LeuArgAspValTrpAspTrpValCysThrIleLeuThrAspPheLysAsnTrpLeuThr	2000	Db	7361	GGAGATGCCCTTCAACAGCTGGCCATTAAAGTCTTCTTGGCCAGCCCCCCCCCAAGCGCGAT	7420
Db	6281	CTCCGGATGTGGGACTGGGTTCGACCATCTAACAGACTTTAAAAAATTGGCTGACC	6340	QY	2361	SerGlyLeuSerThrGlyAlaGlyAlaAlaAspSerGlySerGlnThrProProAspGlu	2380
QY	2001	SerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTyrLys	2020	Db	7421	TCAGGCTTTTCCACGGGGCGGCGCTGCCGATTCGGCAGTCAGACGCTCTCTGATGAG	7480
Db	6341	TCCAAATATTCCAAAGATGCCCGCTCCCTTGTCTCTCTGTCAAAAGGGGTACAG	6400	QY	2381	LeuAlaLeuSerGluThrGlySerIleSerSerMetProProLeuGluGlyGluLeuGly	2400
QY	2021	GlyValTrpAlaGlyThrGlyIleMetThrThrArgCysProCysGlyAlaAsnIleSer	2040	Db	7481	TTGGGCTTTTTCGAGACAGGTTCCATCTCTCCATGCCCTCCCTCGAGGGGAGCTTGA	7540
Db	6401	GGCGTGGGCGCGCACTGGCATCATGACACACACGCTGCTCTTGGCGCGCAATATCTCT	6460	QY	2401	AspProAspLeuGluProGluGlnValGluProGlnProProGlnGlyValAla	2420
QY	2041	GlyAsnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTrp	2060	Db	7541	GATCCAGACTGGAGCTGAGCAGGTAGAGCCCCCAACCCCCCCCCCAGGGGGGGTGGCA	7600
Db	6461	GGCAATGTCGCTTGGGCTCCATGAGAAATCACGGGGCTTAAGACCTGATGAATATCTGG	6520	QY	2421	AlaProGlySerAspSerGlySerTrpSerThrCysSerGluGluAspAspSerValVal	2440
QY	2061	GlnGlyThrPhePheProLeuAsnCysTyrThrGluGlyGlnCysValProLysProAlaPro	2080	Db	7601	GCTCCGGCTCGGACTCGGGGTCTTGGTCTACTTGTCTCCGAGGAGGACGACTCCGCTG	7660
Db	6521	CAGGGGACCTTTCTCATCAATGTTACACGGAGGGCCAGTGGCGGAAACCCCGGCCA	6580	QY	2441	CysCysSerMetSerTyrSerTrpThrGlyAlaLeuIleThrProCysSerProGluGlu	2460
QY	2081	AsnPheLysValAlaIleTrpArgValAlaAlaSerGluTyrAlaGluValThrGlnHis	2100	Db	7661	TGCTGCTCATGTACTACTCTGCGAGGGGTCTTAATAACTCTTGTAGTCCCGAAGAG	7720
Db	6581	AACTTTAAGGTGCCCATCTGAGGGTGGCGGCTCAGAGTACGCGGAGGTGACGCGAGCAC	6640				

ORIGIN	
Alignment Scores:	
Pred. NO.:	0
Score:	16120.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
Matches:	9711
Conservative:	3033
Mismatches:	0
Indels:	0

DB:	6	Gaps:	0
US-09-980-559-2 (1-3033) x AX057395 (1-9711)			
QY	1	MetSerThrAsnProLysProGlnArgLysThrIleAspGlnProGln	20
DB	341	ATGAGCAAAATCTTAACCTCAAGAAACCAAAAGAACCAACCAACCTGCGCCCAA	400
QY	21	AspValLysPheProGlyGlyGlyGlnIleValGlyValTyrLeuLeuProArgArg	40
DB	401	GACGTTAAGTTTCGGCGCGCGCCAGATCGTTGGCGGAGTATACTTGTTCGGCGCAGG	460
QY	41	GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly	60
DB	461	GGCCCCAGGTGGGTGTGCGCGGACAAAGAAAGATTCGAGGCGGTCCAGCCACGCTGA	520
QY	61	ArgArgGlnProIleProLysAspArgArgSerThrGlyLysSerTrpGlyLysProGly	80
DB	521	AGCGCCAGGCCCATCCCTAAGATCGCGCTCCACTGGCAANTCCTGGGGAACACGGA	580
QY	81	TyrProTrpProLeuTyrGlyAsnGluGlyLeuGlyTrpAlaGlyTrpLeuLeuSerPro	100
DB	581	TACCCCTGGCCCTATACGGGAATGAGGACTCGGCTGGCAGAGTGGCTCTGTGCCCC	640
QY	101	ArgGlySerArgProSerTrpGlyProAsnAspProArgHisArgSerArgAsnValGly	120
DB	641	CGAGGTTCCTGCTCCCTCTGGGGCCCCAATGACCCCGGCATAGGTTCGCGCAACGTGGT	700
QY	121	LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProValVal	140
DB	701	ARGGTCAATCATACCTTAACGTGGCTTTGGCGACCTCATGGGGTATACCTCCCTGTGCTG	760
QY	141	GlyAlaProLeuGlyGlyValAlaArgAlaLeuAlaHisGlyValArgValLeuGluAsp	160
DB	761	GGGCGCCCGCTCGCGCGCTGCGCAGAGCTCTCGCGCATCGCTGAGAGTCTTGGAGAC	820
QY	161	GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla	180
DB	821	GGGGTTAAATTTGCAACAGGGAATTAACCGGTGCTCTCTTTCTATCTTTCTTCTTGGCC	880
QY	181	LeuLeuSerCysIleThrThrProValSerAlaAlaGluValLysAsnIleSerThrGly	200
DB	881	CTGCTGTCCTGCAATACCAACCCCGCTCTCGCTGCGGAAGTGAAGAACATCAGTACCGGC	940
QY	201	TyrMetValThrAsnAspCysThrAsnAspSerIleThrTrpGlnLeuGlnAlaAlaVal	220
DB	941	TACATGGTACTAACGACTGCACCAATGACAGCATTTACCTGGCAGCTCCAGGCTGTGTC	1000
QY	221	LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTrpIle	240
DB	1001	CTCCAGTCCCGGGTGGTCCCGTGGCGAAGTGGGAATGCATCTCAGTCTGGATA	1060
QY	241	ProValSerProAsnValAlaValGlnArgProGlyAlaLeuThrGlnGlyLeuArgThr	260
DB	1061	CCGGTCTCACCGAATGTGGCGGTGCGACGCGCCCGCGCCCTCACGCGGGGCTTGGCGAGC	1120
QY	261	HisIleAspMetValValMetSerAlaThrLeuCysSerAlaLeuTyrValGlyAspLeu	280
DB	1121	CACATCGACATGGTGTGATGTCCGCCACGCTCTGCTCTGCCCTCTACGTGGGGACCTC	1180
QY	281	CysGlyGlyValMetLeuAlaAlaGlnMetPheIleValSerProGlnHisIleTrpPhe	300
DB	1181	TGCGGTGGGTGATGCTCGCAGCCCAATGTTTCAATGCTCTCGCGCAGCACCACCTGGTTT	1240
QY	301	ValGlnAspCysAsnCysSerIleTyrProGlyThrIleThrGlyHisArgMetAlaTrp	320
DB	1241	GTCCAAAGACTGCAATTTGCTCCATCTACCTGTGTACCATCATCTGGACACCGCATGGATG	1300
QY	321	AspMetMetMetAsnTrpSerProThrAlaThrMetIleLeuAlaTyrAlaMetArgVal	340
DB	1301	GACATGATGATGAACGTGGTCCGCCACGGCTACCATGATCTTGGCGTACGGATGCTGTC	1360
QY	341	ProGluValIleIleAspIleIleSerGlyAlaHisIleTrpGlyValMetPheGlyLeuAla	360
DB	1361	CCCGAGGTCAATTATAGACATCATTAGCGGGCTCATTTGGCGCTCATTTGGCTTGGCC	1420
QY	361	TyrPheSerMetGlnGlyAlaTrpAlaLysValValIleLeuLeuAlaAlaGly	380
DB	1421	TACTTCTCTATGAGGAGCGTGGGGAAGTCTGTGTATCTCTCTGTGTGGCGCGGG	1480
QY	381	ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr	400
DB	1481	GTGACGCGCGCACCATCATCTGTGTGGGGTTCGCGCGCAGACACCGCGCGCTCACCC	1540
QY	401	SerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTrp	420
DB	1541	AGCTTATTTGACATGGGCCCCAGCGCAAAATCCAGTCTGTTAACCAATGCGAGCTGG	1600
QY	421	HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer	440
DB	1601	CACATCAACCGCACCCCTTGAATGCAATGACTCTCTTGACACCCGGCTTATCGGCTCT	1660
QY	441	LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg	460
DB	1661	CTGTTCTACACCCACAGCTTCAACTCTGTCAGGATGTCGCCAGCGCATGTCGCGCTGCCG	1720
QY	461	SerIleGluAlaPheArgValGlyTrpGlyAlaLeuGlnTyrGluAspAsnValThrAsn	480
DB	1721	AGTATCGAGGCTTCCGGGTGGATGGCGCGCTTGCATATATGAGGATAATGTACCAAT	1780
QY	481	ProGluAspMetArgProTyrCysTrpHisTyrProProArgGlnCysGlyValValSer	500
DB	1781	CCAGAGGATGATGAGACCTTATGCTGGCACTACCCCAAGGCGATGTGGCGTGGCTCC	1840
QY	501	AlaLysThrValCysGlyProValTyrCysPheThrProSerProValValValGlyThr	520
DB	1841	GCGAAGACTGTGTGTGGCCAGTGTACTGTCTTTCACCCCGCAGCCAGTGTGTGGCGACG	1900
QY	521	ThrAspArgLeuGlyAlaProThrTyrThrTrpGlyGluAsnGluThrAspValPheLeu	540
DB	1901	ACCGACAGGCTTGGAGCGCCACTTACAGCTGGGGGAGAAATGAGACAGATGCTTCTCTA	1960
QY	541	LeuAsnSerThrArgProLeuGlySerTrpPheGlyCysThrTrpMetAsnSerSer	560
DB	1961	TTGAAAGGACACTCGACACCGCTGGGGTCTGCTTGGCTGCGACGAGTGAATCTTCT	2020
QY	561	GlyTyrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer	580
DB	2021	GGCTACACCAAGACTTGGCGCGCACCCCTGCGCTACTAGAGTCACTTCAACGCCAGC	2080
QY	581	ThrAspLeuLeuCysProThrAspCysPheArgLysHisProAspThrThrTyrLeuLys	600
DB	2081	ACGGACCTGTGTGCGCCACCGACTGTTTTAGGAAGCATCTCGATACCACTTACCTCAA	2140
QY	601	CysGlySerGlyProTrpLeuThrProArgCysLeuIleAspTyrProTyrArgLeuTrp	620
DB	2141	TGCGGCTCTGGGCGCTTGGCTCACGCCAAGTGGCTGATGCTGCTTACCCCTACAGGCTGTG	2200
QY	621	HisTyrProCysThrValAsnTyrThrIlePheLeuValLeuArgMetTyrValGlyVal	640
DB	2201	CATTACCCCTGACAGTAACTATACCATCTTCAAATTAAGATGATGTGTGGAGGGGTT	2260
QY	641	GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp	660
DB	2261	GAGCACAGGCTCACGCGCTGCATGCAATTTCTCTGTTGGGATGCTGTTGCAACTTGGAGGAC	2320
QY	661	ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGluTrpAlaIleLeuPro	680
DB	2321	AGAGACAGAAAGTCAACTGTCTCTTTGTTGCACTCCACACGGAATGGGCCATTTTACCT	2380
QY	681	CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisGlnAsnIle	700
DB	2381	TGCTCTTACTCGAGCTGCGCGCTTGTGCTGCTGCTTCTTCTCACTCCACCAACATC	2440
QY	701	ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrLysTyrIleValArgTrp	720

2441 GTGAGCTACAAATTCATGTATGCTTATCACCTGCTCCCTCAAAAATACATCGTCCGATGG 2500
 721 GluTrpValIleLeuLeuPheLeuLeuAlaAspAlaArgValCysAlaCysLeuTrp 740
 2501 GAGTGGGTAAATCTCTTATTCCTGCTCTTAGCGAGCGCCAGGGTTTGCCTGCTTATGG 2560
 741 MetLeuIleLeuLeuGlyGlnAlaGluAlaLeuGluLysLeuValIleLeuHisAla 760
 2561 ATGCTCATCTTGTGGCCAGGCGGAAGCAGCTAGAGAGCTGTCTATCTGACGCT 2620
 761 AlaSerAlaAlaSerCysAsnGlyPheLeuTrpPheValIlePhePheValAlaAlaTrp 780
 2621 GCGAGCGCAGCTAGCTGCAATGGCTTCTATATTTTGTCTATCTTTTGTGGCTGCTGG 2680
 781 TyrIleLysGlyArgValValProLeuAlaThrTyrSerLeuThrGlyLeuTrpSerPhe 800
 2681 TACATCAAGGGTGGGTAGTCCCTTATGCTACCTATTCCTCACTGGCTGTGGTCTT 2740
 801 SerLeuLeuLeuAlaLeuProGlnGlnAlaTyrAlaTyrAspAlaSerValHisGly 820
 2741 AGCTACTGCTCTAGCAATGGCCCAACAGGCTTATGCTTATGACGCACTGTGCAATGGC 2800
 821 GlnIleGlyAlaAlaLeuValMetIleThrLeuPheThrLeuThrProGlyTyrLys 840
 2801 CAGATAGGAGCGGCTCTGCTGGTAAATGATCACTCTCTTACTCTCACCCCGGGTATAAG 2860
 841 ThrLeuLeuSerArgPheLeuTrpTrpLeuCysTyrLeuLeuThrLeuGlyGluAlaMet 860
 2861 ACCCTTCTACCGCGGTTTTGTGGTGTGTGCTATCTTCTGACCCCTGGGGGAAAGTATG 2920
 861 ValGlnGluTrpAlaProMetGlnValArgGlyGlyArgAspGlyIleIleIleTrpAla 880
 2921 GTCCAGAGTGGGACCACTATGACAGTGGCGGTGGCTGATGCAATCATATGGGCC 2980
 881 ValAlaIlePheTrpProGlyValValPheAspIleThrLysTrpLeuLeuAlaValLeu 900
 2981 GTCCCATATTCACCCAGGTGGTGTGTGACATACCAAGTGGCTTCTTGGCGGTGCTT 3040
 901 GlyProAlaTyrLeuLeuLysGlyAlaLeuThrArgValProTyrPheValArgAlaHis 920
 3041 GGGCTGCTTACCTCTCAAAAGGTGCTTTGACGCGGTGCTGCTTCTGTCAGGGCTCAC 3100
 921 AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyGlyArgTyrValGlnMet 940
 3101 GCTCTACTAGGATGTGCACATGGCAAGGCACTTCGCGGGGGCAGGTACGTCCAGATG 3160
 941 AlaLeuLeuAlaLeuGlyA-gfTrpThrGlyThrTyrIleTyrAspHisLeuThrProMet 960
 3161 GCGCTACTAGCCCTTGGCAGGTGACCTGGCACTTACATCTATGACCACTCACCCCTATG 3220
 961 SerAspTrpAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIleIlePhe 980
 3221 TCGGATTTGGGTGCTAGTGGCTCGCGGACCTTGGCGGTGCTTGGCTTATCATCTTC 3280
 981 SerProMetGluLysLysValIleValTrpGlyValAlaGluThrAlaAlaCysGlyAspIle 1000
 3281 AGTCCGATGAGAAAGTCAATGTCTGGGAGCGGAGACAGCTGCTTGTGGGACAT 3340
 1001 LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp 1020
 3341 TTACACGGACTTCCCGGTCCGCGGACCTTGGTGGGAGGTCTCTTGGCCAGCTGAT 3400
 1021 GlyTyrThrSerLysGlyTrpSerLeuLeuAlaProIleThrAlaTyrAlaGlnGlnThr 1040
 3401 GGCTATACCTCAAGGGGTGGAGTCTTCTCGCCCCCATCATCTGCTTACGCCACAGACA 3460
 1041 ArgGlyLeuLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGluGlnAla 1060
 3461 CGTGGCTTTTGGCACCATAGTGGTGAACATGACGGGGCGGCACAGACAGAGCGCT 3520
 1061 GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly 1080
 3521 GGGGAAATTCAGTCTCTGTCACAGTCACTCAGTCTCTTCTCGGAAACATCATCTCGGG 3580

1081 ValLeuTrpThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro 1100
 3581 GTTITGTGACGCTGTACCATCGAGCTGGCAACAGACTCTTGGCGGCTCACGGGTCGG 3640
 1101 ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTyrProSerProGly 1120
 3641 GTCAGCAGATGCTACTCCAGTGTCTGAGGGGAGCTTAGTAGGGTGGCCAGCCCTGGG 3700
 1121 ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArgAsn 1140
 3701 ACTAATATCTTTGGAGCCGTCACGTGTGGAGCGGTGCACTGTACCTGTGTGACCGGAAC 3760
 1141 AlaAspValIleProAlaAlaArgArgGlyAspLysArgGlyAlaLeuLeuSerProArg 1160
 3761 GCTGATGCTATCCCGGCTCGAAGACCGGGGAGCNAACGGGAGCGCTACTCTCCCCGAGA 3820
 1161 ProLeuSerThrLeuLysGlySerSerGlyGlyProValLeuLeuCysProArgGlyHisAla 1180
 3821 CCTCTTTCCACCTTGAAGGGTCTCTCAGGAGGCCGCTGCTATGCCCGGAGGCGCACGCT 3880
 1181 ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle 1200
 3881 GTCGGAGTCTTCGGGAGCTGTGCTCTCGGGGCGTGGCTAAAGTCCATAGATTTCATC 3940
 1201 ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro 1220
 3941 CCGCTTGACAGACTGCACATCTGCACGCTCCCCCACCCTTTAGTGCACACAGCACCA 4000
 1221 ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLys 1240
 4001 CCTGCTGTCCCCAGACCTATCAGTCTGGGTACTTGCATGTCGCCGACTGGCAGTGAAG 4060
 1241 SerThrLysValProValAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnPro 1260
 4061 AGCACCAAGTCTCTGTCATATGCTGCTCAGGGGTATAAAGTGTAGTGTCTTAAATCCC 4120
 1261 SerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerLysAlaHisGlyIleLeuAsnPro 1280
 4121 TCAGTGGCTGCCACCTCGGGGTTTGGGGGCTACTGTCTTAAGGCACATGGCATCATCCC 4180
 1281 AsnIleArgThrGlyValArgThrValThrGlyAlaProIleThrTyrSerThrTyr 1300
 4181 AACATTAGGACTGGAGTGGAGTGTGACGACCGGGGCGCCCATCATCATCATATAT 4240
 1301 GlyLysPheLeuAlaAspGlyGlyCysAlaGlyAlaTyrAspIleIleIleCysAsp 1320
 4241 GGCATAATCTCTCGCCGATGGGGCTGTGGGGGCGGCTTACGACATCATATATGTAT 4300
 1321 GluCysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGlnAla 1340
 4301 GAATGCTATGCTGGACTTACACCATCTTGGCATCGGAACAGTCTCTTGTATCAAGCA 4360
 1341 GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProGlySerValThr 1360
 4361 GAGACAGTGGGTGAGACTAACTGTGCTGGCTACAGTACGCGCCCTGGGTGAGTCA 4420
 1361 ThrProHisProAsnIleGluGluValAlaLeuGlyGlnGluGlyGluIleProPheTyr 1380
 4421 ACCCCCCACCCCAACATAGAGGAGTGGCCCTTGGGCGAGGGCGAGATCCCTCTCTAT 4480
 1381 GlyArgAlaIleProLeuSerTyrIleLysGlyGlyArgHisLeuIlePheCysHisSer 1400
 4481 GGGAGGCGGATTCCTCTGCTTACATCAAGGAGGAGAGACATCTGATCTCTTCCCATCA 4540
 1401 LysLysLysCysAspGluLeuAlaAlaLeuArgGlyMetGlyLeuAsnSerValAla 1420
 4541 AAGAAAAAGTGTGACGAGCTCGCGGCGCCCTTCCGGGTATGGGTGTAACCTCAGTGGCA 4600
 1421 TyrTyrArgGlyLeuAspValSerValIleProThrGlnGlyAspValValValAla 1440
 4601 TACTACAGAGGGTGGACGCTCTCCGTAATCAACCACTCAGGAGACGCTAGTGGTGGTGGC 4660

QY	1441	ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnVal	1460	Db	5741	CTCACCAGTCCGCTGTCAACAGCACCACTATCTCTCAACATTTTGGGGGCTGGCTA	5800
Db	4661	ACCGAGCCCTCATGACAGGGTACTTGGGGACTTTGACTCGGTGATCGACTCAACGTA	4720	QY	1821	AlaSerGlnIleAlaProProAlaGlyAlaThrGlyPheValValSerGlyLeuValGly	1840
QY	1461	AlaValThrGlnValValAspPheSerLeuAspProThrPheThrIleThrGlnIle	1480	Db	5801	GCATCCAAATTTGCACACCCCGGGGGCCACTTGGCTTCTGTGTAGTGGCTTAGTGGGA	5860
Db	4721	GGCGTCACTCAAGTTGTAGACTTCAGTTTAGACCCACATTCACCAATAACACACAGATT	4780	QY	1841	AlaAlaValGlySerIleGlyLeuGlyLysValLeuValAspIleLeuAlaGlyTyrGly	1860
QY	1481	ValProGlnAspAlaValSerArgSerGlnArgArgGlyArgGlyArgGlyArgLeu	1500	Db	5861	GCTGCCGTAGGCAGTATAGCTTAGGTAAAGTGTAGTGCACATCTTGGCAGGATGGT	5920
Db	4781	GTCCCTCAAGACGCTGTCTACGTAGCCAGCGCGGGTGCACAGGTAGGGGAAGACTG	4840	QY	1861	AlaGlyIleSerGlyAlaLeuValAlaPheLysIleMetSerGlyGlyLysProSerMet	1880
QY	1501	GlyIleTyrArgTyrValSerThrGlyLysArgAlaSerGlyMetPheAspSerValVal	1520	Db	5921	GGCGGCAATTTGGGGGCTCTCGTGCATTCAGATCATCTTGGCGAGAGCCCTCCATG	5980
Db	4841	GGCAATTTATAGGTATGTTTCCATGTGTAGCGAGCCCTCAGGAATGTTTACAGTGTAGT	4900	QY	1881	GluAspValValAsnLeuLeuProGlyIleLeuSerProGlyValAlaLeuValValGlyVal	1900
QY	1521	LeuCysGluCysTyrAspAlaGlyAlaAlaTyrTyrGluLeuThrProSerGluThrThr	1540	Db	5981	GAGGATGTCTCACTTGTCTGTGATTTCTTCCGGGTGCTTGTGTAGTGGAGTCT	6040
Db	4901	CTCTGTAGTGTCTACGACGAGGGGCGCATGTGTATGAGTCTACCATCTGGAGACACC	4960	QY	1901	IleCysAlaAlaIleLeuArgArgHisValGlyProGlyGlyGlyAlaValGlnTyrMet	1920
QY	1541	ValArgLeuArgAlaTyrPheAsnThrProGlyLeuProValCysGlnAspHisLeuGlu	1560	Db	6041	ATCTGCGCGCCATTTCTGCGCCGACACACGTGGGACCGGGGGAGGCGCGCTCCAATGGATG	6100
Db	4961	GTCCAGGCTCAGGCGGTATTTCAACACGCGCGGTTTGCCTGTGTGCCAAGACCATCTT	5020	QY	1921	AsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaProThrHisTyrValThr	1940
QY	1561	PheTyrGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThr	1580	Db	6101	AATAGACTCATTTGCTTCTCCAGAGGAAATCATCTGCGCCGCCACCCACTACGTGACG	6160
Db	5021	TTTTGGGAGGCAATTTTACCGGCTCACACATAGATGCCCACTTCTCTTCCCAACA	5080	QY	1941	GluSerAspAlaSerGlnArgValThrGlnLeuLeuGlySerLeuThrIleThrSerLeu	1960
QY	1581	LysGlnSerGlyGluAsnPheAlaTyrLeuThrAlaTyrGlnAlaThrValCysAlaArg	1600	Db	6161	GAGTGGATGCGTGTGCGAGGTGTGACCCAACTACTTGGCTCTCTTACCATAACACGCTG	6220
Db	5081	AAGCAATCGGGGAAATTTTCGATATCTTAACAGCCCTTACAGGCTCAGTGTGCGTAGG	5140	QY	1961	LeuArgArgLeuHisAsnTyrIleThrGluAspCysProIleProCysGlyGlySerTyr	1980
QY	1601	AlaLysAlaProProSerTyrAspValMetTyrLysCysLeuThrArgLeuLysPro	1620	Db	6221	CTCAGAGACTCCACACTGGATTACTGAGGACTGCCCATCCCATGCGCGGCTCGTGG	6280
Db	5141	GCCAAGACCCCGCCCGCTCTGCGAGCTCATGTGGAGTTTGTGACTGACATCAAGCCC	5200	QY	1981	LeuArgAspValTyrAspTyrValCysThrIleLeuThrAspPheLysAsnTyrLeuThr	2000
QY	1621	ThrLeuValGlyProThrProLeuLeuTyrArgLeuGlySerValThrAsnGluValThr	1640	Db	6281	CTCCGCGATGTGTGGGACTTGGCCATCTAACACAGACTTTAAAAATTTGGCTGACC	6340
Db	5201	ACACTGTGGGCCCCACACCTCTCTGTACCGCTTGGGCTCTGTATACCAACAGGTCACC	5260	QY	2001	SerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTyrLys	2020
QY	1641	LeuThrHisProValThrLysTyrIleAlaThrCysMetGlnAlaAspLeuGluValMet	1660	-Db	6341	TCCAAATTAATCCAAAGATGCCCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	6400
Db	5261	CTCACACATCCCGTGACAAATACATCGCCACCTGCTGACCAAGCCGCTTGAAGTCTG	5320	QY	2021	GlyValTyrAlaGlyThrGlyIleMetThrThrArgCysProCysGlyAlaAsnIleSer	2040
QY	1661	ThrSerThrTyrValLeuAlaGlyGlyValLeuAlaAlaValAlaAlaTyrCysLeuAla	1680	Db	6401	GGCGTGTGGCGCGCACCTGGCATCATGACACACAGGTGTCTTGGCGGCCAATATCTCT	6460
Db	5321	ACCAGCACATGGGTCTTGGCAGGGGAGTCTTGGCGGCGCTTCCCGGATTCCTTGGCG	5380	QY	2041	GlyAsnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTyr	2060
QY	1681	ThrGlyCysValCysIleIleGlyArgLeuHisIleAsnGlnArgAlaValAlaPro	1700	Db	6461	GGCAATGTCCGCTTGGGCTCCATGAGATCAGCGGGCTTAAGACCTGCATGAATATCTGG	6520
Db	5381	ACCGGCTGTCTTGTGATCATCGGCGCTTGCACATTAACCCAGCGAGCGCTTGGCGCG	5440	QY	2061	GlnGlyThrPheProIleAsnCysTyrThrGluGlnCysValProLysProAlaPro	2080
QY	1701	AspLysGluValLeuTyrGluAlaPheAspGluMetGluCysAlaSerArgAlaAla	1720	Db	6521	CAGGGAGCTTTCTTATCAATTTGTACAGGAGGGCCAGTGGCTGCCGAAACCCCGGCCA	6580
Db	5441	GACAAGAGGTCTCTATAGGCTTTGTATGAGATGGAGGAATGTCCCTCTAGGGCGGCT	5500	QY	2081	AsnPheLysValAlaIleTyrArgValAlaAlaSerGluTyrAlaGluValThrGlnHis	2100
QY	1721	LeuIleGluGlyGlnArgIleAlaGluMetLeuLysSerLysIleGlnGlyLeuLeu	1740	Db	6581	AACTTTAAGTGTCCCATCTGGAGGGTGGCGGCTCAGAGTACCGGAGGTGACGACGAC	6640
Db	5501	CTCATTTAAGAGGGGCGAGATAGCGAGATGCTGAAGTCCAAAGATCCAAAGCTTATG	5560	QY	2101	GlySerTyrHisTyrIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeu	2120
QY	1741	GlnGlnAlaSerLysGlnAlaGlnAspIleGlnProThrValGlnAlaSerTyrProLys	1760	Db	6641	GGGTATACCACTACATAACAGGACTCCACCTGATTAATCTGAAGTCCCTGCCAATA	6700
Db	5561	CAGCAAGCTTCCAAACCAAGCTCAAGACATACACCCACCTGTGCGAGGCTTCATGGCCCAAG	5620	QY	2121	ProSerProGluPhePheSerTyrValAspGlyValGlnIleHisArgPheAlaProThr	2140
QY	1761	ValGluGlnPheTyrAlaLysHisMetTyrAsnPheIleSerGlyIleGlnTyrLeuAla	1780	Db	6701	CCCTCTCCGAGTCTCTTCTTGGGGTGGAGGAGTGCAGATCCATAGGTTTGGCCCCACA	6760
Db	5621	GTAGAACAATTTCTGGGCGCAACACATGTGGAACTTCAATTAGCGGCACTCAATACCTCGCA	5680	QY	2141	ProLysProPhePheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal	2160
QY	1781	GlyLeuSerThrLeuProGlyAsnProAlaValAlaSerMetMetAlaPheSerAlaAla	1800	Db	6761	CCGAGCCGTTTTTTCGGGATGAGGTCTCTGCTTGGGCTTAATTCATTTGTCGTC	6820
Db	5681	GGACTATCAACATGCCAGGGAACCTTGCAGTAGCTTCCATGATGGCGTTCAGTCCGCC	5740	QY	2161	GlySerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThr	2180
QY	1801	LeuThrSerProLeuSerThrSerThrThrIleLeuLeuAsnIleLeuGlyGlyTyrLeu	1820				

6821 GGCTCCAGCTTCCTTCGACCTCGAACCCGACACAGACGTATTCATGTCCTAACA 6880
 2181 AspProSerHisIleThrAlaGluThrAlaAlaAraArgLeuAlaArgGlySerProPro 2200
 6881 GATCCATCTCATATACGGCGGAGACTGCGAGCGCGGCTTTAGCGCGGGGGTACCCCCA 6940
 2201 SerGluAlaSerSerAlaSerGlnLeuSerAlaProSerIleuAlaThrCysThr 2220
 6941 TCCGAGGCAAGCTCTCCGGGAGCGAGCTATCGGACACCATCGCTGGAGCCACTGCACC 7000
 2221 ThrHisGlyLysAlaTyrAspValAspMetValAspAlaAsnLeuPheMetGlyGlyAsp 2240
 7001 ACCACGGCAAGCCCTATGATGGACATGGTGGATGCTAAACCTGTTCATGGGGCGCAT 7060
 2241 ValThrArgIleGluSerGlySerLysValValValLeuAspSerLeuAspProMetVal 2260
 7061 GTGACTCGGATAGACTCTGGGTCTCAAGTGGTCTGTGGACTCTCTCGACCCCAATGGTC 7120
 2261 GluGluArgSerAspLeuGluProSerIleProSerGluTyrMetLeuProLysLysArg 2280
 7121 GAAGAAAGGAGCGACCTTGAGCCTTCGATACCATCAGATACATGCTCCCAAGAGAGG 7180
 2281 PheProAlaLeuProAlaTyrAlaArgProAspTyrAsnProProLeuValGluSer 2300
 7181 TTCCACACAGCTTTACCGGCTCGGCACGCGCTGATTACAAACCCACCGCTTGTGGAATCG 7240
 2301 TrpLysArgProAspTyrGlnProAlaThrValAlaGlyCysAlaLeuProProArg 2320
 7241 TGGAAAGGCGCAGATTACCAACCGGCGCTGTTGGCGGCTGTCTCTCCCTCTCTCTAGG 7300
 2321 LysThrProThrProProAlaArgArgArgArgValGlyLeuSerGluAspSerIle 2340
 7301 AAAACCCGAGCGCTCCCAACAGAGCGCGCGAGCAGTGGCGCTTAAGTGGAGCTCCATA 7360
 2341 GlyAspAlaLeuGlnLeuAlaIleLysSerPheGlyGlnProProSerGlyAsp 2360
 7361 GGAGATGCCCTTCAACAGCTGGCCATTAACTCTCTTTGGCGAGCGCGCGCGAT 7420
 2361 SerGlyLeuSerThrGlyAlaGlyAlaAlaAspSerGlySerGlnThrProProAspGlu 2380
 7421 TCAGAGCTTTCCACAGCGGCGCGCTGCGCATTCGCGCAGTCAGACGCGCTCTGTAGTAG 7480
 2381 LeuAlaLeuSerGluThrGlySerIleSerSerMetProProLeuGluGlyLeuGly 2400
 7481 TTGGCCCTTTCCGAGACAGGTTCACTCTTCATGCCCGCTTCGAGGGGAGCTTGA 7540
 2401 AspProAspLeuGluProGluGlnValGluProGlnProProProGlnGlyValAla 2420
 7541 GATCCAGACCTGGAGCTGAGCAGGTAGAGCCCAACCCCGCGCGGGGGTGGCA 7600
 2421 AlaProGlySerAspSerGlySerTrpSerThrCysSerGluGluAspSerValVal 2440
 7601 GCTCCCGGCTCGGACTCGGGGCTGTGCTTACTTGTCTCGAGGAGGACGACTCCGCTGTG 7660
 2441 CysCysSerMetSerTrpTrpTrpGlyValAlaLeuIleThrProCysSerProGluGlu 2460
 7661 TGCTGCTCCATGTCATCTCTGACCGGGGCTCTAATAACTCTTGTAGTCCGAGAGAG 7720
 2461 GluLysLeuProIleAsnProLeuSerAsnSerLeuLeuArgTyrHisAsnLysValTyr 2480
 7721 GAGAGTTACCGATTAAACCCCTTGAGCAACTCCCTGTTGGATATCAACAAGGTGTAC 7780
 2481 CysThrThrThrLysSerAlaSerLeuArgAlaLysLysValThrPheAspArgMetGln 2500
 7781 TGTACCAACAAGAGCGGCTCACTAAGGGCTTAAAGAGTAACTTTTGTATGATGATGCAA 7840
 2501 ValLeuAspSerTyrTyrAspSerValLeuLysAspIleLysLeuAlaAlaSerLysVal 2520
 7841 GTGCTCGACTCTCTACTAGACTCAGTCTTAAAGACATTAAGCTAGCGGCTCCAGGTC 7900
 2521 ThrAlaArgLeuLeuThrMetGluAlaCysGlnLeuThrProProHisSerAlaArg 2540
 7901 ACCGCAAGGCTCTCTACCATGGAGGAGGCTTGCCAGTTAACCACCCCACTTCTGCAAGA 7960

2541 SerLysTyrGlyPheGlyValAlaLysGluValArgSerLeuSerGlyValArgAlaValAsnHis 2560
 7961 TCTAATATGGTGGGCTAAGGAGGTTCGACAGCTTGTTCGGGAGGCGCGTAAACCAC 8020
 2561 IleLysSerValTrpLysAspLeuLeuGluAspSerGluThrProIleProThrThrIle 2580
 8021 ATCAAGTCCGTGGAGGACCTCTCGAGGACTCAGAACCAACCAATTCCTCAACCAAT 8080
 2581 MetAlaLysAsnGluValPheCysValAspProThrLysGlyGlyLysAlaAlaAsp 2600
 8081 ATGGCCAAAATAGGTGCTTCGCGGTGACCCCAAGAGGGGGCAAGAAAGCAGCTCGC 8140
 2601 LeuIleValTyrProAspLeuGlyValArgValCysGluLysMetAlaLeuTyrAspIle 2620
 8141 CTTATGCTTTACCTGACCTCGCGCTCAGGCTTCGCGAGAGATGGCCCTTTATGACAT 8200
 2621 ThrGlnLysLeuProGlnAlaValMetGlyAlaSerTyrGlyPheGlnTyrSerProAla 2640
 8201 ACACAAAACCTTCTCAGCGGTGATGGGGCTTCTTATGGATTCAGTATTCCTCCCGCT 8260
 2641 GlnArgValGluPheLeuLeuLysAlaTyrAlaGluLysLysAspProMetGlyPheSer 2660
 8261 CAGCGGTAGAGTTCTCTTGAAGCATGGCGGAAAGAGAGCCCTATGGGTTTTTCG 8320
 2661 TyrAspThrArgCysPheAspSerThrValThrGluArgAspIleArgThrGluGluSer 2680
 8321 TATGATACCCGATGCTTTCACCTCAGCTCAGGAGAGACATCAGGACTGAGGAGTCC 8380
 2681 IleTyrArgAlaCysSerLeuProGluGluAlaHisThrAlaIleHisSerLeuThrGlu 2700
 8381 ATATATCGGCGCTCTCTTCGCGAGGAGGCGCACCTGCATACACTGCCTAACTGAG 8440
 2701 ArgLeuTyrValGlyGlyProMetPheAsnSerLysGlyGlnThrCysGlyTyrArgArg 2720
 8441 AGACTTTAGCTGGGAGGCGCTATGTTCAACAGCAGAGGCGCAACCTGCGGGTACAGCGCT 8500
 2721 CysArgAlaSerGlyValLeuThrThrSerMetGlyAsnThrIleThrCysTyrValLys 2740
 8501 TGCGCGCCAGCGGGGTGCTCACCCTAGCATGGGGAACACCATCATCATGCTACGTGAAA 8560
 2741 AlaLeuAlaAlaCysLysAlaAlaGlyIleIleAlaProThrMetLeuValCysGlyAsp 2760
 8561 GCCTTAGCGGCTTGAAGCTCAGGAGTAATCGCGCCCAACATGCTGTGTTGCGGCGAT 8620
 2761 AspLeuValValIleSerGluSerGlnGlyThrGluGluAspGluArgAsnLeuArgAla 2780
 8621 GACTTGGTGTATCTCAGAAAGCCAGGGGACCGAGAGGACGAGCGGAACCTTGAGAGCC 8680
 2781 PheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProArgProGluTyr 2800
 8681 TTCAGGAGCTATGACAGGATTTCTGCGCTCTCTGCTGAGCCCGCCAGACCGAGTAT 8740
 2801 AspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaLeuGlyProGlnGly 2820
 8741 GATCTGAGCTGATAACATCTTGTCTCTCAATGTCTGTGTGCGCTGGGCGCCCAAGGC 8800
 2821 ArgArgArgTyrTyrLeuThrArgAspProThrThrProIleAlaArgAlaAlaTrpGlu 2840
 8801 CGCGCAGAGTACTTACCTGACAGAGCCCTTACCATCTCCCAATCGCGCGGCTGCTGGAA 8860
 2841 ThrValArgHisSerProValAsnSerTrpLeuGlyAsnIleIleGlnTyrAlaProThr 2860
 8861 ACAGTTAGACATCTCCCTGTCAATTCATGGTGGGAAACATCATCCAGTACGCCCGGACC 8920
 2861 IleTrpAlaArgMetValLeuMetThrHisPhePheSerIleLeuMetAlaGlnAspThr 2880
 8921 ATATGGGCTCGCATGGTCTCTGATGACACACTTCTTCTCCATTCATGCTCAGACAGC 8980
 2881 LeuAspGlnAsnLeuAsnPheGluMetTyrGlyValValTyrSerValSerProLeuAsp 2900
 8981 CTGACCAAGAACCTCACTTTGAGATGTAGCGAGCGGTGTACTCCGTGAGTCCCTTGGAC 9040

QY 2901 LeuProhlailelleGluAargLeuHisGlyLeuAspAlaPheSerLeuHisThrTyThr 2920
 Db 9041 CTCACGCTATAATTGAAGGTTACATGGCTTTCAGCGCTTTCTCTGCACATACACT 9100
 QY 2921 ProHisGluLeuThrArgValAlaSerAlaLeuArgGlyLeuGlyAlaProProLeuArg 2940
 Db 9101 CCCACGACTGACAGGGTGGCTTCAGCCCTCAGAAACTTGGGGCGCCACCCCTCAGA 9160
 QY 2941 AlaTrpLysSerArgAlaArgAlaValArgAlaSerLeuLeuSerArgGlyGlyArgAla 2960
 Db 9161 GCGTGAAGAGCGCGGACGTCAGTCAGTCAGGCGTCCCTCATCTCCCGTGGGGGAGAGCG 9220
 QY 2961 AlaValCysGlyArgTyLeuPheAsnTrpAlaValLysThrLysLeuLysLeuThrPro 2980
 Db 9221 GCGCTTTGCGCTGCATATCTCTTCAATTGGCGCGTGAAGCAAGCTCAAACTCACTCCA 9280
 QY 2981 LeuProGluAlaArgLeuLeuAspLeuSerSerTrpPheThrValGlyAlaGlyGly 3000
 Db 9281 TTCCCGGAAGCGCGCTCTCGAATTTATCCAGCTGGTTCAACCGTCGCGCGCGCGGGCG 9340
 QY 3001 AspIleTyHisSerValSerArgAlaArgProArgLeuLeuLeuPheGlyLeuLeu 3020
 Db 9341 GACATTATCACAGCGTGTGCGTGGCGGACCGCTTATGCTCTTTGGCCTACTCCTA 9400
 QY 3021 LeuPheValGlyValGlyLeuPheLeuLeuProAlaArg 3033
 Db 9401 CTTTGTAGGGTAGGCGCTTCTTCTACTCTCCCGCTCGG 9439
 RESULT 4
 AF177036
 LOCUS
 DEFINITION Hepatitis C virus clone pJ6CF, complete genome. VRL 05-OCT-1999
 ACCESSION AF177036
 VERSION AF177036.1 GI:6010579
 KEYWORDS
 SOURCE
 ORGANISM
 Hepatitis C virus
 Hepatitis C virus
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 Hepacivirus.
 1 (bases 1 to 9711)
 Yanagi, M., Purcell, R.H., Emerson, S.U. and Bukh, J.
 Hepatitis C virus: an infectious molecular clone of a second major
 genotype (2a) and lack of viability of intertypic 1a and 2a
 chimeras
 Virology 262 (1), 250-263 (1999)
 MEDLINE 99420396
 PUBMED 10489358
 REFERENCE 2 (bases 1 to 9711)
 Bukh, J.
 Direct Submission
 Submitted (11-AUG-1999) Hepatitis Viruses Section, Laboratory of
 Infectious Diseases, National Institute of Allergy and Infectious
 Diseases, National Institutes of Health, Building 7, Room 201,
 7-Center Dr., Bethesda, MD 20892-0740, USA
 Location/Qualifiers
 1. .9711
 /organism="Hepatitis C virus"
 /mol_type="genomic RNA"
 /strain="HC-J6CH"
 /db_xref="taxon:11103"
 /clone="pJ6CF"
 /notes="infectious cDNA"
 genotype: 2a"
 341..9442
 /codon_start=1
 /product="polyprotein"
 /protein_id="AA011178.1"
 /db_xref="GI:6010580"
 /translation="MSTNPKPKTKNTNRRPQDVKFGGGQIVGVVLLPRGPRLL
 GVPAKTSERSQPRRRQPIPKDRSTGKSGKPGYPPLVNEGLGWAGWLLSPRG
 SRPSWGNPFRHRSNRGVKIDLTGCFADLMGYIPVAGPLGVGARAHAHGVRLVD
 GNVFATGNLPGCSFSIFLLALLSCITTPVSAAEVKNISVYMWYNDCTNDSITWQLQA
 AVLHVPFCVPCKEVGNASQCWIIPVSPNVAVQRGALTOGLRTHIDMVMSATLCSALY

VGDLCGVMLAQMFTIVSPQHWFVQDCNCSYIPGTITGHRMAWMMNWSPTATMIL
 AYAMRVEPIIDIIISGAHGMVNFGLAYFSMQGAWAKVVVILLAAAGDARTHTYTGSA
 AOTGRLTSLDFMGPRQTLQVNTGSMHINRTALNDSLTGFTASTIPTYTHFSNS
 GCPERMSACRSIEAFRGVGAQLOYEDNTNPMPCYCHYPRQCGCTWNSAGVTCGPV
 YCFPSVUVYTTDRLGAPTYTNGENEDVFLNSTRPPLGSGFSGCTWNSAGVTCYC
 GAPCRTRADFNASTDLCLPTDCFRKHPDTTYLKGSGPWLPRCLIDTPYRLWHYPC
 YVNTYIFKIRMTVGGVHEHLTAACNFTRGDRCNLEDRDRSQSLPSTHSTTEWAILPCS
 YSDLPALSTGLLHLHQNIVDVQFMYGLSPALTKYIVRMEWVILLFLLADACACILM
 MTLILQAARALEKIVILHAASAACNGFLYFVIFVAAAYIKRIVVPLATVSLTGLM
 SFSLLLALUPOQAYIVDASVGOIGAAALVMTLFTLTTPGYKTLRSFLWATLYLLTL
 GEMVQEWAPPMQVRGGRDGIIVAVAI FYGVVEDITKWLAVLGPAYLLKGLTRVP
 FYRAHALLRMCTMARHLAGGRVQNALALGRWGTYYIYDHLTPVNDWAASGLRDLA
 VAVEPIIFSPMEKKVIVMGAETAACGDIILHGLPVSARLREVLTPADGTYSGWLSL
 APITAYAQOTRGLGTYIVSMGTGRDKEQAGEIQLSTVQSPIGTSISGLVATVTHG
 AGNKTLAGSGPVTQMYSSAEGDLVGAEPGPGTKSLEPCTCGAVDLYLVTRNADYIP
 RRGDKRGALLSPRLSTLKGSSGGPVLCPRGHVGVPRAAVCSRGVAKSIDPIVET
 LDIIVRSPTFSDNSPPAVPQYGVYLAFTGSGSKTKVFNAYAAQGYKVLVNFVS
 ADLGFGLTSLKAHGINPRTIGRTVITGAPITYSTYKFGADGGCAGYDIIICD
 ECHAVDSTTILGIGTILDOAETAGVRLATATATPGSVTTTPHNTPEEVALGQGEIP
 FYGRAIPLSYIKGGRHILIFCHSKKCKDELAALRGMLNSVAYIRGLDSVITQGDV
 VVATDALMTGYTGDFSDVIDCNVATVQVDFSLDPTFTTQIIVPQDAVSRSORGR
 TGRGLGIYRVYSTGERASCMFSDVLCYDAGAAWYELTPSETTVRLRAYPNTPL
 PYQDHFLEWNAVFTGLTHIDAFPLSOTKQSGENFAYLTAYQATVCARAKAPPNDV
 MMKLTLPRLTIVGPTPLLYRLGVSNTVEVLTHTPVTKYIATCMQADLEWMTSTWLAG
 GVLAAYVAYCLATGCVCIIGRLHINORAVAPDKVLYEAFDEMEECAARAALIEEQ
 RIAEMLKSKIQGLLOQASQAQDIQPTVQASWPQVQFWAKHMMNFISGTQYLAGLST
 LIPNPVAMSAFSAALTSPLSTSTILLNLGLWLASOIAIPAGATGFWVGLVGA
 VGSIGLGVLDILAGYGAGISGALVAFKIMSGEKPSMEDVNNLLPGLISPGALVGV
 ICAAILRHVGPGEAGVQMNRLIAFASGRNHVAPHYVTHYTESDASORVTLGLSLIT
 SLRRLHNNITDCPIPCGSLRDVDMVCTILTDTDKNLTSLKLPKMPGLFPFVSCQ
 KGYKPAVGTGIMTTRPCGANSIGNVRILGSMRTITGPKTCMNIWQSTFPINCYTEQC
 VQPPAPFVKAIRVAAASEAYEVTHQSYHYITGLTDLNLPVQCLPSPFPFVSDVG
 QLRFPATPKPFPRDEVSFCVGLNSFVGSOLPCDPEPTDVMMLTDPDHTIESTA
 ARRLARSGSPSSASSASOLAPSLRATCTTHGKAVDMDVMDLPMGDDVTRIESG
 KYVLDSDLPVTEERSDLSPSIPEYMLPKBFPALPAWADPNPVLVESWKRDPY
 QPATVAGCALPPPKPTPPRRRTVSLGSESDIGDALQOLAIKSGQPPPSGDSGLS
 TGAAGASGSPDPDELALSETGSISSMPLEGLGDPDLEPEQVPPQPPQGGVAAAP
 GSDGSWSSTSEEDSVVCSMSYNTGALITPCSPBEKLPINPLNSILLRYTHNYKY
 CTTTGSASLRARKKVTFRMDVLDYSYDVKDIKLAASKVARTALLTMEACQLTPHIS
 ARSKYGFGAKEVRSLSGRAVNIKSVKDKLLEDSETPITPTIMAKNEVFCVDPTKGGK
 KAARLIVPDLGVRCERKALYDITQKLPOAVNGASYGQYSPAQRVEFLKWAEEK
 DPMGYSYDFRCSDFSTVERDIRTESIYRACSLPSEAHAIHSLRLYVGGFVNSK
 GTCGYRCRASGVLVTSMGNTITCYKALAAKAAAGIIAPTMLVCGDDLVVISEYQ
 TEEDERNLRAFTEAMTRYSPADPPRDEYDELEITSCSSNVSVAGLQDQRRYYLST
 DPTPTIARAWEYTVRHSVNSWLGNTIOVAPTIWARMVMTHTFSLTMAQDQTDONLN
 FEMYGANVSVPLDLPAILERLHGLDASFSLHTYTHELTVASALKLGPPLRAKWS
 RARAVASLISRGRAAVCGRYLFWAVTKLKLPLPEARLLDLSSWTFVAGGGDI
 YHSVSRARPRLLLFGLLLLFVGVGLFLLPAR"

BASE COUNT 1987 a 2853 c 2668 g 2203 t

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 9711
 Score: 16120.00 Matches: 3033
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 14 Gaps: 0

US-09-980-559-2 (1-3033) x AF177036 (1-9711)

QY 1 MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGln 20
 Db 341 ATGAGCACAAATCTTAACCTCAAGAAAAACCAAGAAAAACCAACCGTCCCCACAA 400
 QY 21 AspValLysPheProGlyGlyGlyGlnIleValGlyGlyValTyrlLeuLeuProArgArg 40
 Db 401 GACGTTAAGTTTCGGGGCGCGCCAGATCGTTGGCGGAGTATACCTTGTTCGCCGCGAG 460
 QY 41 GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly 60
 Db 461 GGCCCCAGGTTGGTGTGCGCGCAGCAAGAAAGACTTCGGAGGGGTGCCAGCCACGTGGA 520

[illegible]

Db	2681	TACATCAAGGGTCGGGTAGTCCCTTAGCTACCTATTCCCTCACTGCGCTGTGCTCTTT	2740
Qy	801	SerLeuLeuLeuLeuAlaLeuProGlnGlnAlaTyrAlaTyrAspAlaSerValHisGly	820
Db	2741	AGCTACTGCTCTAGCATTTGCCCAACAGGCTTAGTCTATGACGCACTGTGTCATGGC	2800
Qy	821	GlnileGlyAlaAlaLeuLeuValMetIleThrLeuPheThrLeuThrProGlyTyrLys	840
Db	2801	CAGATAGGAGCGGCTCTGCTGGTAAATGATCACTCTCTTTACTCTCACCCCGGGTATAG	2860
Qy	841	ThrLeuLeuSerArgPheLeuTyrThrLeuCysTyrLeuLeuThrLeuGlyGluAlaMet	860
Db	2861	ACCTCTCTAGCCGGTTTGTGGTGTGTGTCTATCTCTGACCCCTGGGGGAAGCTATG	2920
Qy	861	ValGlnGluTrpAlaProMetGlnValArgGlyGlyArgAspGlyIleIleTrpAla	880
Db	2921	GTCCAGAGTGGGACCACTATGACAGTGGCGGTGGCGGTGATGGCATATATGGGCC	2980
Qy	881	ValAlaIlePheTyrProGlyValValPheAspIleThrLysTrpLeuLeuAlaValLeu	900
Db	2981	GTCCGCATATTTCTACCCAGGTGGTGTGTGTGACATACCAAGTGGCTCTTGGCGGTGCTT	3040
Qy	901	GlyProAlaTyrLeuLeuLysGlyAlaLeuThrArgValProTyrPheValArgAlaHis	920
Db	3041	GGGCTGTCTTACCTCTAAAGGTGCTTTGACGCGGTGCGTACTTCTGTCAGGGCTCAC	3100
Qy	921	AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyGlyArgTyrValGlnMet	940
Db	3101	GCTCTACTGAGGATGTGCACCATGGCAGGCACTCTCGGGGGGAGGTACGTCCAGATG	3160
Qy	941	AlaLeuLeuAlaLeuGlyArgTyrTrpThrGlyThrTyrIleTyrAspHisLeuThrProMet	960
Db	3161	GGCTACTAGCCCTTGGCAGGTGGACTGGCCACTTACATCTATGACCACTCACCCCTATG	3220
Qy	961	SerAspTrpAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIleIlePhe	980
Db	3221	TCGGATTTGGGCTGCTAGTGGCCCTCGGGAGCTTGGCGGTGCGCTTGGAGCTATCATCTTC	3280
Qy	981	SerProMetGluLysLysValIleValTrpGlyValaGluThrAlaAlaCysGlyAspIle	1000
Db	3281	AGTCCGATGAGAGAAAGTCATTTCTGGGAGCGAGACAGCTGCTTGTGGGGACATTT	3340
Qy	1001	LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp	1020
Db	3341	TTACACGCACTTCCCGGTGCTCGCCGACTTGGTGGGAGGTCTCTTGGCCCGAGCTGAT	3400
Qy	1021	GlyTyrThrSerLysGlyTrpSerLeuLeuAlaProIleThrAlaTyrAlaGlnGlnThr	1040
Db	3401	GGCTATACCTTCAAGGGGTGGAGTCTTCTCGCCCCCATCACTGCTTACGCCCAAGCAGACA	3460
Qy	1041	ArgGlyLeuLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGluGlnAla	1060
Db	3461	CGTGGCCCTTTGGGCACCATAGTGGTGAGCATGACGGGGGCGGACAGACAGACAGGCT	3520
Qy	1061	GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly	1080
Db	3521	GGGGAAATTCAGTCTGCTCCACAGTCACTCAGTCTCTCTCTCGAACATCCATCTCGGG	3580
Qy	1081	ValLeuTrpThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro	1100
Db	3581	GTTTTGTGGACTGTCTACCATGGAGCTGGCAACAGACTCTGGCCGCTCACGGGGTCCG	3640
Qy	1101	ValThrClnMetTyrSerSerAlaGluGlyAspLeuValGlyTrpProSerProProGly	1120
Db	3641	GTCAAGCAGATGTACTCCAGTGTGAGGGGAGCTTAGTAGGGTGGCCAGCCCTCTGGG	3700
Qy	1121	ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArgAsn	1140
Db	3701	ACTAAATCTTTGAGCGGTGCACGTGTGGAGCGGTGACCTGTACTGTGTCAAGCGGAC	3760
Qy	1141	AlaAspValIleProAlaArgArgArgGlyAspLysArgGlyAlaLeuLeuSerProArg	1160
Db	3761	GCTGATGTATCCCGGCTCGAAGACGCGGGGCAAAACAGGGGAGCGGTACTCTCCCGGAGA	3820
Qy	1161	ProLeuSerThrLeuLysGlySerSerGlyGlyProValLeuCysProArgGlyHisAla	1180
Db	3821	CCTCTTTCCACCTTGAAGGGTCTCTCAGAGGGCCCGTGTCTATGCCCGGGCCACGCT	3880
Qy	1181	ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle	1200
Db	3881	GTGGAGTCTTCGGGCAGCTGTGTCTCTCGGGGCGTGGCTAAGTCCATAGATTTTCATC	3940
Qy	1201	ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro	1220
Db	3941	CCGTTTGACACACTCCACATCTGACCGGTCCCCCACCCTTTAGTGACCAACACACCA	4000
Qy	1221	ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaAlaProThrGlySerGlyLys	1240
Db	4001	CCTGCTGTGCCCCAGACCTATCAGGTGGGTCTTGTCTTAAGGCACATGGCATCAATCCC	4060
Qy	1241	SerThrLysValProValAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnPro	1260
Db	4061	AGCACCAAGTTCTGTCTGCATATGTCTCAGGGGTATAAAGTGTAGTGTCTTAATCCC	4120
Qy	1261	SerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerLysAlaHisGlyIleAsnPro	1280
Db	4121	TCAGTGGCTGCCACCCTGGGGTTTGGGGGTACTTGTCTTAAGGCACATGGCATCAATCCC	4180
Qy	1281	AsnIleArgThrGlyValArgThrValThrThrGlyValaProIleThrTyrSerThrTyr	1300
Db	4181	ACATTTAGAGCTGGAGTCAAGACTGTGACGACGGGGGCCCATCATCTACTCCACATAT	4240
Qy	1301	GlyLysPheLeuAlaAspGlyGlyCysAlaGlyGlyAlaTyrAspIleIleIleCysAsp	1320
Db	4241	GGCAAAATTCCTCGCGATGGGGCTGTGGCGGGCGGCTTACGACATCATATATGTAT	4300
Qy	1321	GluCysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGlnAla	1340
Db	4301	GAATGCCATTCGCGTGGACTCTACCACTCTTGGCATCGGAACAGTCTCTTGTATCAAGCA	4360
Qy	1341	GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProGlySerValThr	1360
Db	4361	GAGACAGCTGGGTCAAGACTTACTGTCTGGCTTACAGCTACGCCCTTGGTCACTGACA	4420
Qy	1361	ThrProHisProAsnIleGluGluValAlaLeuGlyGlnGluGlyGluIleProPheTyr	1380
Db	4421	ACCCCCACCCCAACATAGAGAGGTGGCCCTTGGGCAGAGGGGAGAGTCCCTCTCTAT	4480
Qy	1381	GlyArgAlaIleProLeuSerTyrIleLysGlyGlyArgHisLeuIlePheCysHisSer	1400
Db	4481	GGAGGGGCGATTTCCCTGTCTTACATCAAGGGGAGGAAGACATCTGATCTTCTGCCATTC	4540
Qy	1401	LysLysLysCysAspGluLeuAlaAlaLeuArgGlyMetGlyLeuAsnSerValAla	1420
Db	4541	AGAAAAAGTGTGACGAGCTCGGGCGGCCCTTGGGGGTATGGGCTTGAACCTCAGTGGCA	4600
Qy	1421	TyrTyrArgGlyLeuAspValSerValIleProThrGlnGlnGlyAspValValValAla	1440
Db	4601	TACTACAGAGGTGGACGCTCTCCGTAAATACCACTCAGGAGAGCGTAGTGGTCTCGCC	4660
Qy	1441	ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnVal	1460
Db	4661	ACCGACGCCCTCATGACGGGTATCTGGGGGACTTTTGACTCCGTGATCGACTGCAACGTA	4720
Qy	1461	AlaValThrGlnValValAspPheSerLeuAspProThrPheThrIleThrThrGlnIle	1480
Db	4721	GGGTCACTCAAGTTGTAGACTTTCAGTTTAGACCCCCACATTCACCAATACCAACAGATT	4780
Qy	1481	ValProGlnAspAlaValSerArgSerGlnArgArgGlyArgThrGlyArgGlyArgLeu	1500
Db	4781	GTCCCTCAAGACGCTGTCTCAGCTAGCAGCGCGGGGTGCGACGGGTAGGGGAAGACTG	4840
Qy	1501	GlyIleTyrArgTyrValSerThrGlyGluArgAlaSerGlyMetPheAspSerValVal	1520
Db	4841	GGCATTTATAGGTATGTTTCCACTGTGTGAGCGGCGCTCAGGAATGTTTGACAGTGTAGTG	4900

QY 1521 LeuCyseGluCysTyrAspAlaGlyAlaAlaTyrTyrGluLeuThrProSerGluThrThr 1540
 DB 4901 CTCTGTAGTGTCTACGACGCGGGCGCGATGTATGAGCTACACCATCGGAGACCACC 4960
 QY 1541 ValArgLeuArgAlaTyrPheAenThrProGlyLeuProValCysGlnAseHisLeuGlu 1560
 DB 4961 GTCAGGCTCAGGCGGTATTTCAACACGCGCGTTTCCCTGTGTGCCAAGACCATCTTGAG 5020
 QY 1561 PheTyrGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThr 1580
 DB 5021 TTTTGGGAGGAGTGTTCACCGCGCTCACACATAGATGCCACCTTCCTTTTCCCAACA 5080
 QY 1581 LysGlnSerGlyGluAenPheAlaTyrLeuThrAlaTyrGlnAlaThrValCysAlaArg 1600
 DB 5081 AACCAATCGGGGGAATTTCCGATATCTTAACAGCCTACACGCTACAGTGTGGCTAGG 5140
 QY 1601 AlaLysAlaProProSerTyrAspValMetTrpLysCysLeuThrArgLeuLysPro 1620
 DB 5141 GCCAAGAGCCCCCGCTCTGGGAGCTCATGTGGAAGTGTGTGACTCGACTCAAGCCC 5200
 QY 1621 ThrLeuValGlyProThrProLeuLeuTyrArgLeuGlySerValThrAenGluValThr 1640
 DB 5201 ACATCGTGGGCCCCACACCTCTCTGTACCGCTTGGGCTCTGTATCAACAGAGGTCAAC 5260
 QY 1641 LeuThrHisProValThrLysTyrIleAlaThrCysMetGlnAlaAspLeuGluValMet 1660
 DB 5261 CTCACACATCCCGTACGAAATACATCGCCACCTGCATGACCGACCTTGAGGTCAATG 5320
 QY 1661 ThrSerThrTyrValLeuAlaGlyGlyValLeuAlaAlaValAlaAlaTyrCysLeuAla 1680
 DB 5321 ACCAGCACATGGGTCTTGGCAGGGGAGTCTTGGCGCGCTCGCGGTATTCCTGGCG 5380
 QY 1681 ThrGlyCysValCysIleIleGlyArgLeuHisIleAenGlnArgAlaValAlaPro 1700
 DB 5381 ACCGGGTGTGTTCATCATCGCGCGCTTGCACATTAACACGAGCGCGTCTGTGGCGG 5440
 QY 1701 AspLysGluValLeuTyrGluAlaPheAspGluMetGluGluCysAlaSerArgAlaAla 1720
 DB 5441 GACAGAGGTCTCTATGAGGCTTTTATGAGATGAGAGAGATGTGCTCTAGGGCGGT 5500
 QY 1721 LeuIleGluGlyGlnArgIleAlaGluMetLeuLysSerLysIleGlnGlyLeuLeu 1740
 DB 5501 CTCATTCAAGAGGGGCGGATAGCCGAGATGCTGAAGTCCAAGATCCAAGGCTTATTG 5560
 QY 1741 GlnGlnAlaSerLysGlnAlaGlnAspIleGlnProThrValGlnAlaSerTyrProLys 1760
 DB 5561 CAGCAAGCTTCCAACCAAGCTCAAGACATCAACCCACTGTGAGGCTTCATGGCCCAAG 5620
 QY 1761 ValGluGlnPheTyrAlaLysHisMetTyrAsnPhelIleSerGlyIleGlnTyrLeuAla 1780
 DB 5621 GTAGAACAAATCTGGGCGCAACACATGTGGNACTTCAATTAGCGGCATCCATACCTGCA 5680
 QY 1781 GlyLeuSerThrLeuProGlyAsnProAlaValAlaSerMetMetAlaPheSerAlaAla 1800
 DB 5681 GGACTATCAACACTGCCAGGGAACCTCGAGTAGCTTCCATGATGGGCTTCAGTGGCGCC 5740
 QY 1801 LeuThrSerProLeuSerThrSerThrThrIleLeuLeuAenIleLeuGlyGlyTyrLeu 1820
 DB 5741 CTCACAGTCCGCTGTCAACAGCACCACTATCTCTCAACATTTTGGGGGGCTGGCTA 5800
 QY 1821 AlaSerGlnIleAlaProProAlaGlyAlaThrGlyPheValValSerGlyLeuValGly 1840
 DB 5801 GCATCCCAAAATGTGACCAACCGCGGGGCGCACTGGCTGTGTGTGTCAGTGGCTTAGTGGGA 5860
 QY 1841 AlaAlaValGlySerIleGlyLeuGlyLysValLeuValAspIleLeuAlaGlyTyrGly 1860
 DB 5861 GCTCCGTAGGCACTATAGGCTTAGGTAAAGTGTAGTGGACATCTCTGGCAGGGTATGGT 5920
 QY 1861 AlaGlyIleSerGlyAlaLeuValAlaPheLysIleMetSerGlyGlyLysProSerMet 1880
 DB 5921 GCGGGCATTTTGGGGGCTCTCGTCGCATTTCAAGATCATGTCTGCGGAGAAGCCCTCCATG 5980
 QY 1881 GluAspValValAenLeuLeuProGlyIleLeuSerProGlyAlaLeuValValGlyVal 1900

DB 5981 GAGGATGCTCAACTGCTGCTGGAATTCGTCTCCGGGTGCTCTGTAGTGGAGTC 6040
 QY 1901 IleCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTyrMet 1920
 DB 6041 ATCTCGGGCGCATTTCTGCGCGACACGTTGGGACCGGGGGAAGGCGCGCTCAATGGATG 6100
 QY 1921 AsnArgLeuIleAlaPheAlaSerArgGlyAenHisValAlaProThrHisTyrValThr 1940
 DB 6101 ANTAGACTCATTTGCTTTCAGAGGAATACAGTCCGCCCCACCCACCTACGTGACG 6160
 QY 1941 GluSerAspAlaSerGlnArgValThrGlnLeuLeuGlySerLeuThrIleThrSerLeu 1960
 DB 6161 GAGTCGATGCGTCGCGCGTGTGACCCCAACTACTTGGCTCCCTTACCATTAACACGCTG 6220
 QY 1961 LeuArgArgLeuHisAsnTrpIleThrGluAspCysProIleProCysGlyGlySerTyr 1980
 DB 6221 CTCAGAAGACTCCCAACTGGATTACTGAGGACTGCCCATCCCATGCGCGGCTCGTGG 6280
 QY 1981 LeuArgAspValTyrAspTyrValCysThrIleLeuThrAspPhelLysAsnTyrLeuThr 2000
 DB 6281 CTCGCGCATGCTGGGACTGGGTTCACCATCTCTAACAGACTTTAAAAATTTGGCTGACC 6340
 QY 2001 SerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTyrLys 2020
 DB 6341 TCCAAATATTCCCAAGATGCCGCGCTCCCTTGTCTCTGTCAAAAGGGGTACAAG 6400
 QY 2021 GlyValTyrAlaGlyThrGlyIleMetThrThrArgCysProCysGlyAlaAenIleSer 2040
 DB 6401 GCGGTGTGGCGCGCATCGGCATCATGACACACGCGTGTCTTGGCGCGCAATATCTCT 6460
 QY 2041 GlyAsnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAenIleTyr 2060
 DB 6461 GGCATGTCCGTGGGTCTTCATGAGATCACGGGCGCTTAAGACCTGATGATATCTGG 6520
 QY 2061 GlnGlyThrPheProIleAsnCysTyrThrGlyGlnCysValProLysProAlaPro 2080
 DB 6521 CAGGGACCTTCTCTATCAATTTGTACACGGGGCGAGTGGCTGCGGAAACCGCGCCA 6580
 QY 2081 AsnPhelLysValAlaIleTyrArgValAlaAlaSerGluTyrAlaGluValThrGlnHis 2100
 DB 6581 AACTTTAAGGTGCCCATCTGGAGGGTGGCGGCTCAGAGTACGCGGAGGTGACGACAC 6540
 QY 2101 GlySerTyrHisTyrIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeu 2120
 DB 6641 GGGTCATACCATACATACAGGACTCACCATGATACTTGAAGTCCCTCTGCCAACA 6700
 QY 2121 ProSerProGluPhePheSerTyrValAspGlyValGlnIleHisArgPheAlaProThr 2140
 DB 6701 CCTCTCTCCGAGTCTCTTCTGGGTGGACGGAGTGCAGATCCATAGGTTTGGCCCCACA 6760
 QY 2141 ProLysProPhePheArgAspGluValSerPheCysValGlyLeuAenSerPheValVal 2160
 DB 6761 CCGAAGCCGTTTTTCGGGATGAGGTCTGCTTCTGGTGGGCTTAATTCATTTGCTGTC 6820
 QY 2161 GlySerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThr 2180
 DB 6821 GGGTCCAGGCTTCTTGGACCCCTGAACCCGACACAGAGCTATGTATGTCATGCTAACCA 6880
 QY 2181 AspProSerHisIleThrAlaGluThrAlaAlaArgArgLeuAlaArgGlySerProPro 2200
 DB 6881 GATCCATCTCATATACGCGCGAGACTGACGCGCGGCTTTAGCGGGGGGTCAACCCCA 6940
 QY 2201 SerGluAlaSerSerAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThr 2220
 DB 6941 TCCGAGCAAGCTCTCCGCGAGCAGCTATCGGCACCATCTGCTGGGAGCCACTGACCC 7000
 QY 2221 ThrHisGlyLysAlaTyrAspValAspMetValAspAlaAenLeuPheMetGlyGlyAsp 2240
 DB 7001 ACCACGCGCAACCTTATGATGTGGACATGTGTGATGCTTAACCTGTTTCATGGGGGCGAT 7060
 QY 2241 ValThrArgIleGluSerGlySerLysValValValLeuAspSerLeuAspProMetVal 2260

Copied from 09540843 on 05/19/2004

2981	LeuProGluAlaAArgLeuLeuAspLeuSerSerTrpPheThrValClyAlaGlyGly	3000
Qy		
9281	TTGCGGGAAGCGCGCTCTCGATTTATCCAGCTGGTCCACGTCGGCGCCGGCGGGGC	9340
Ddb		
3001	AspIleTyrHisSerValSerArgAlaAArgProAArgLeuLeuPheGlyLeuLeuLeu	3020
Qy		
9341	GACATTATCACAGGTGTGCGGTGCCGACCCCGCTATTGCTCTTTGGCTACTCTTA	9400
Ddb		

3021	LeuPheValGlyValGlyLeuPheLeuLeuProAlaArg	3033
Qy		
9401	CTTTTGTAGGGTAGGCCCTTTTCTACTCCCGCTCG	9439
Db		

RESULT 5	
HPCPOLP	HPCPOLP
LOCUS	9589 bp RNA linear VRL 20-JUN-1998
DEFINITION	Hepatitis C virus genomic RNA for polyprotein, complete cds.
ACCESSION	D00944
VERSION	D00944.1 GI:221650
KEYWORDS	polyprotein.
SOURCE	Hepatitis C virus
ORGANISM	Hepatitis C virus
	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
	Hepadnavirus.

REFERENCE
1. (bases 1 to 9599)
Miyakawa, Y., Okada, S., Sugiyama, Y., Kurai, K., Izuka, H., Machida, A., Okamoto, H., and Mayumi, M.
Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved and divergent regions
J. Gen. Virol. 72 (Pt 11), 2697-2704 (1991)

MEDLINE
PUBMED
PUBMED
1658196
2 (sites)
Han, J.H. and Houghton, M.
Group specific sequences and conserved secondary structures at the
3' end of HCV genome and its implication for viral replication
Nucleic Acids Res. 20 (13), 3520 (1992)
JOURNAL

MEDLINE	92335016
PUBMED	1321416
REFERENCE	3 (sites)
AUTHORS	Hotta, H., Doi, H., Hayashi, T., Purwanta, M., Soemarto, W., Mizokami, M., Ohba, K. and Homma, M.
TITLE	Analysis of the core and E1 envelope region sequences of a novel variant of hepatitis C virus obtained in Indonesia
JOURNAL	Arch. Virol. 136 (1-2), 53-62 (1994)
MEDLINE	94270990
PUBMED	7545932
COMMENT	These data kindly submitted in computer readable form by: Hiroaki

```

tax:
FEATURES             source
    Location/Qualifiers
        1..9589
            /organism="Hepatitis C virus"
            /mol_type="genomic RNA"
            /isolate="HC-J6"
            /db_xref="taxon:11103"
        341..9442
            /codon_start=1
            /product="polyprotein"
            /protein_id="BAA00792.1"
            /gi="221651"
            /translation="MSTNPKPKKTKNTNRPDQVKFPGGGQIVGGVYLLPRRGPRLL
GVATRKTSERSQPRGRQDIPKDRSTGSKWGKPGYPWPLYGNEGLGWAGWLLSPRG
SPRSWPNDRPHRSNRFTDITLCGFADLMGYIPVVGAPLGCVARALAHGVRLVED
GVNFATNLPFGCSFTFLKALLSCITTPVSAAEKNIISCTGYMTNDCTNDSITWQIQAA
AVLHVPCGVVPCEKVGNTSRCTIPVSPNVAQPGALQTGLRTHIDMVNMSATTCALY
VLDLCCGVMLAOMFTVSPHWHFVQDNCSTVPGTITGRHWADMMNMSPTATMIL

```

AYANRVEVIDIIGGAHGWMPGCLAYF5MQGAKVAVVILLLAAHTGAVDAOHTYVGGST
 AHNARTUTMFSGARQKIQLINTGNSWHINTALNCNDLSHTGTFASLFTYTHSPNS
 CGPERMSACSIEAFRVMGALQYEDNVNTPEDRMPYCMHYPPQCGVSVSASSVCGVP
 YCFTRSPVVTYDRLGALPTTCTDDEFTVLLNSTRPPQGSWFCTGNNSTGYTKTC
 GACPRTRADFNASMDLLCPTDCPKRHDPTTYIKCGSGPMLTPCLDIYDMLWHYPC
 VYNDIPIKIRYMGVGBHRUTAOACNFTRGDNLEDRDSQLSPLLHSTTBWAILPCT
 TSYDPLTGLLHLHONIVDVMYMGSPALTKYIWRWVWVLLPLLADARVACCLIA
 MLILPQGBAALAEKLVILHRAASAASCNGLFYVYVFFVAWYIKGRVPLATYSLTGLW
 SFGULLLLPQQAAYVDA5VHQIGAAULLVLTITLTPGYKTLISRLFWMLWLYLLTL
 AENVQWAPPMQVWRGGRDGI1WAV1PCFQGVEDITKWLVLALGPVLLKGLALTRV
 YVFAHALLRMTVRHLAGRYVQYMWLLALGLRWGTGYIYDHLTFMSPDWAANGRLDA
 VAFVPI1PSMKEKVIYMGABACGDI1LHGLPV5ARLGRVLLGPDAGTY5KGNWLL
 APITAYAAQTRGLLGTIV5WBGDRDKTEQAGEIOVLSTVQSPGLTTISGLVLTVYHG
 ANKTLAGSRGPVTQMTYS5BAGDVLGH9PPSPKSLBECTCGADVLIATVRADNADIPA
 RRGKGALLSPQML5LTKGSSGCPVLCPRGHAVGVFRAACVSRGVAKGIDTPIFVET
 LDIVTRFETSDHNPINPTQYVGYLHAPTGSKGSTKVPVAAQGYKVLVILNPSV
 ADITLFGAYLSKSHGINTPNTQYVTVTTCAPIT5TYTGKFLADGGCAGYVLI1ICD
 ECHAVDSTTILGIVLDOQBATGRVTLVATATPPG5VTTPHNIEBALVQGEQEP
 FYGR4IPL5YIKGGRHLIFCHSKKGLDELAAALRGMLNAVAYRGLDVSIVTQGDV
 VVATADLGYRTGDFDSVDCNVAVTQVDVSLDPTFTITQTVPODASVRSORRG
 TGRGRLGIYVY5TGERASGDMFDSVLCECYDAGAAWVPLTAVRVLRAVFNTPGL
 PVCQDLHFBEAVFTLTHDAHLSQTKSGRGNFAYLTAQATVCARAKAPPSNDV
 MWKCLTRVLTPGFTPLLYRGLSVNVEVTLHPVTKYATCEQDLBEMVST5VWLAG
 GVLAAAYACLATOCVCIIGRLHYNRVAVAPDEKVEYAEFDEBEC5ARAL1BEGQ
 L1BMLKSKIQGLLQKASQAKQDTPQPAVQSPKVEQFWAKHMFNFISGIVQYLAGLST
 RPAENASVMAFSAALTSPLUSTSTILLNTLGGWLASQIAAPPAGATFVSGVLGVAA
 VGSILGKVLVDILAGYGAGISGALVAFKIMSGEKSPK5MEDVNLPLLGLTSPGALVVG
 ICAALLRHVGPGBGAVQWNNRL1APASRGNHVAPTHVYTESDASVOTQLLGSJIT
 SLRLRLNHHEDVCTPI1CQSGWLADNDWCVCTI1LDPKNNLTSKLFPMKPGGLPFSQC
 KYKGVWAGTGMITTRPCGANISGNVRLGSMRITGPCTCMN1WQGTFFINCYTBEGQC
 VPEKPAFKTAIMRAVASEYAEVTHQS9HYITGLTIDNLKVPQCOLPSP5F5VDDGV
 QIHRFADIPKPFDFE5VFCVGLNSFVVSQOLPCDPEDPDVLMTLDPDHSHTAETA
 ARRLRASGPPEAS5SASQASPLRATCTTHGAYDVMVDVMDLFMGDDVTR1BES
 KVVVLSLDLPMVER5DLERS1P5EYMLPKKRFPPALPAWPDYNPLV5SPKRPD
 QPATVAGCALPPEKKTPTPPRRRTVGLSESS1ADALQOLA1K5FGOPPPSGD5GLS
 TGAADASGKTPDELA15TG5ISMPPLLEG5PGDPLD5PEQVELQPPQCGVWTP
 GSGSG5W5TCEED5VSC5MSY5GAL1PCSPEEK5VTL1CPS5EALQVTHNNKY
 CTT5K5SLAKRYTFFORMQDALHYDSVLKD1KLAASKVTLARLL1B5EACOLTPPHS
 KARKYFGKAEVRS15R5C5ANVH1K5WVKDLLED1QTPIPT1IMAKNEVFCVDPTKGGK
 KAARL1YVPDGLGVR5C5RMALYD1TQKLQPAVMGASGYFOY5PAQRVE5FLKMAWEK
 DPMG5YDTRCFDSVTRD1B5E5IYRAC5LPEAHTA1H5LTERLYV5G5PM5NK
 QOTCGYR5CRAS5VLT5MGNT1TCYKALA5ACKAGH1APTMLV5GDDLVJ15ESQ
 TEEDERNLRAF5AM7Y5APGDP5PR5PE5T5DEL1T5SC5NV5VAL5G5QGR5RYLTR
 DPPT1P5AAW5ETVH5SPVNS5WLGN1IQAYT1WAR5VLM1MTPH5L1MAQDQTLQNLN
 FEMYGAV55P5LQDLP5A1IERLHG5DAL5P5TYTH5LTRV5ASLRK5GAPL5RKS
 RARV5RASL15GRG5RA5C5G5R5FL5W5P5K5L5KL5PL5EARLL1D555FTV5GAGGDI
 YV5V55R5APR1L1L5GLL1L5FV5GVL5FL5PAR"

BASE COUNT	1970 a	2820 c	2633 g	2166 t
ORIGIN				
Alignment Scores:				
Pred. No.:	0		Length:	9589
Score:	15830.00		Matches:	2967
Percent Similarity:	98.81%		Conservative:	30
Best Local Similarity:	97.82%		Mismatches:	36
Query Match:	98.20%		Indels:	0
DB:	14		Gaps:	0

US-09-980-559-2 (1-3033) x HPCPOLP (1-9589)

Qy	1	MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgProGln	20
Db	341	ATGACCAAAATCCTAAACCTCAAAAGAAAAACCAAAAGAAACACCAACCGTCGCCACAA	400
Qy	21	AspValLysPheProGlyGlyGlyGlnIleValGlyGlyValTyrlleuProArgArg	40
Db	401	CACGTTAAAGTTTCGGGGCGGGCGGCAGATCGTTGGCGGAGTATCTTGTTCGGCGCAGG	460
Qy	41	GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly	60
Db	461	GGCCCCAGGTTGGGTGTGGCGCGCACAGGAAGACTTCGGAGCGGTCCACGACCACTGGA	520
Qy	61	ArgArgGlnProIleProLysAspArgArgSerThrGlyLysSerTrpGlyLysProGly	80

521	AGCGCCAGCCCATCCCTAAGGATCGCGCTCCACTGGCAAAATCCTGGGGAAACACAGGA	580	1601	CACATCAACCGCACCGCCTGAACTGCAATGACTCTTTCCACACCGCGCTCTCCCGCTCA	1660	
81	TyrProTrrProLeuTyrGlyAsnGluGlyLeuGlyTrrAlaGlyTrrLeuLeuSerPro	100	QY	LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg	460	
581	TACCCCTGGCCCTATACCGGAATGAGGACTCGCTGGCAGAGTGGCTCTCTGTCCTCC	640	DB	CTGTTCTACACCCACAGCTTCACTGCTCAGGATGTCCTCGAAGCATGTCGCGCTCGCC	1720	
101	ArgGlySerArgProSerTrrProAsnAspProArgHisArgSerArgAsnValGly	120	QY	SerIleGluAlaPheArgValGlyTrrGlyAlaLeuGlnTrrGluAspAsnValThrAsn	480	
641	CGAGGTTCCTCCCTCTCTGGGGCCCAATGATGATCCCGGCATAGGTCCCGCAACGTGGT	700	DB	1721	AGTATCAGGCGCTTTCGGGTGGGATGGGGCGCTTACAAATATGAGGACAATGTCCCAAT	1780
121	LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTrrIleProValVal	140	QY	481	ProGluAspMetArgProTrrCysTrrHisTrrProProArgGlnCysGlyValValSer	500
701	AAGGTCAATGATACCTTACGTCGGCTTTCGGGACCTCATGGGGTACATCCCTGTCGA	760	DB	1781	CCAGAGGATATGAGACCGGATTTGCTGGCACTATCCCAAGACAGTGTGGTAGTCTCC	1840
141	GlyAlaProLeuGlyGlyValAlaArgAlaLeuAlaHisGlyValArgValLeuGluAsp	160	QY	501	AlaLysThrValCysGlyProValTrrCysPheThrProSerProValValValGlyThr	520
761	GGCGCCCGCTCGCGCGGCTGCGCAGAGCTCTCGCGCATGCGTGAAGTCTCGAGGAC	820	DB	1841	GGGAGCTCTGTGTGGCCAGTGTACTGTTCACCCCGCCAGCCAGTAGTAGTGGGTACG	1900
161	GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla	180	QY	521	ThrAspArgLeuGlyAlaProThrTrrTrrTrrGlyGluAsnGluThrAspValPheLeu	540
821	GGGGTTAAATTTGCAACAGGGAATTTACCCGGTGTCTCTTTCTATCTTCTTGTGGCC	880	DB	1901	ACCGATAGACTTGGAGCGCCACTTACACGTGGGGGAGAAATGAGACAGATGTCTTCTTA	1960
181	LeuLeuSerCysIleThrThrProValSerAlaAlaGluValLysAsnIleSerThrGly	200	QY	541	LeuAsnSerThrArgProProLeuGlySerTrrPheGlyCysTrrTrrMetAsnSerSer	560
881	CTGCTGTCTTGCATCACCCCGCTCTCCGCTGCGCAAGTGAAGAACATCAGTACCGGC	940	DB	1961	TTGAAACAGCAGCTCGACCCGCGGCTCATGCTTCGGCTGCACGTGGATGAATCCACT	2020
201	TyrMetValThrAsnAspCysThrAsnAspSerIleThrTrrGlnLeuGlnAlaVal	220	QY	561	GlyTrrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer	580
941	TACATGGTGACCAACAGCTGCAATGATAGCATTTACCTGGCAATCCAGGCTGTGTC	1000	DB	2021	GGCTACCAACAGACTTGGCGGCGCACCCCTCGCGCATTTAGAGCTGACTTCAATGCCAGC	2080
221	LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTrrIle	240	QY	581	ThrAspLeuLeuCysProThrAspCysPheArgLysHisPheAspThrTrrTrrLys	600
1001	CTCCAGCTCCCGGGTGGCTCCGTCGCGAAGTGGGAATACATCTCGTCTCGGATA	1060	DB	2081	ATGGACTTGTGTGCCCGCAGGACTGTTTTAGGAAGCATCTGTATACCACTTACATCAA	2140
241	ProValSerProAsnValAlaValGlnArgProGlyAlaLeuThrGlnGlyLeuArgThr	260	QY	601	CysGlySerGlyProTrrLeuThrProArgCysLysLeuIleAspTrrProTrrArgLeuTrr	620
1061	CCGGTCTCACCGAATGTGGCGTGCAGCCAGGCTTCACTGCTCGCACAGCACCACTGGT	1120	DB	2141	TGTGGCTCTGGCCCTGGCTCGCCAGGTGCTGATCGACTACCCCTACAGGCTCTGG	2200
261	HisIleAspMetValValMetSerAlaThrLeuCysSerAlaLeuTrrValGlyAspLeu	280	QY	621	HisTrrProCysThrValAsnTrrIlePheLysIleArgMetTrrValGlyGlyVal	640
1121	CATTTGACATGTTGTGATGTCGGCCAGCTCTGCTCGCTCTTTACGTGGGGACCTC	1180	DB	2201	CATTACCTCGACAGTTAACTATACCATCTTCAAAATAGGATGTATGTGGGGGGGTC	2260
281	CysGlyGlyValMetLeuAlaAlaGlnMetPheIleValSerProGlnHisTrrPhe	300	QY	641	GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp	660
1181	TGCGGTGGGTGATGCTTGAGCCAGATGTTCACTGCTCGCACAGCACCACTGGT	1240	DB	2261	GAGCACAGGCTCACGCTCGGTGCAATTTCACTCGTGGGGATGCTGTGCAACTTGGAGGAC	2320
301	ValGlnAspCysAsnCysSerIleTrrProGlyThrIleThrGlyHisArgMetAlaTrr	320	QY	661	ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGluTrrPheIleLeuPro	680
1241	GTGCAAGACTGCAATTTGCTCATCTACCTCGGTACCATCATCGACACCGCATGGCGTG	1300	DB	2321	AGAGACAGAAAGTCAACTGTCTCTGCTGCTCCACACGAGGTGGGCGCATTTTACCT	2380
321	AspMetMetMetAsnTrrSerProThrAlaThrMetIleLeuAlaTrrAlaMetArgVal	340	QY	681	CysSerTrrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisIleuHisIleuHis	700
1301	GACATGATGATGAATGCTGCGCCAGCGCTTACCATGATCTGCGGTGACGCGATGGCGTC	1360	DB	2381	TGCATTTACTCGGAGCTGCGCCCTGTCGATGCTGCTTTCACCTCCACCAAAACATC	2440
341	ProGluValIleIleAspIleIleSerGlyAlaHisTrrGlyValMetPheGlyLeuAla	360	QY	701	ValAspValGlnPheMetTrrGlyLeuSerProAlaLeuThrLysTrrIleValArgTrr	720
1361	CCCGAGGTCAATAGACATCATTTGGGGGCTCATTTGGGGGCTCATTTGGGCTTAGCC	1420	DB	2441	GTGGAGTGCATTTCAATGATGCGCTTACCTGCTCTCACAANAATACATCGTCCGATGG	2500
361	TyrPheSerMetGlnGlyAlaTrrAlaLysValValIleLeuLeuLeuAlaGly	380	QY	721	GluTrrValIleLeuLeuPheLeuLeuAlaAspAlaArgValCysAlaCysLeuTrr	740
1421	TACTTCTATGAGGAGCGTGGGCAAAAGTCTGTTGTCATTCTTTTGTGGCGCGCGG	1480	DB	2501	GAGTGGGTAGTACTCTTATTCCTGCTTAGCGGAGCGCCAGGCTTGGCTGTATGG	2560
381	ValAspAlaArgThrHisThrValGlyLysAlaAlaGlnThrThrGlyArgLeuThr	400	QY	741	MetLeuIleLeuGlyGlnAlaGluAlaLeuGluLysLeuValIleLeuHisAla	760
1481	GTGAGCGCGCAACCATACCTTGGGGTTCCTACCGCGCATTAACCCAGGACCTCCACC	1540	DB	2561	ATGCTCATCTTGTGGCCAGGCGCCAGGAGCAGTACAGAAAGTGGTCTGTTCGACGCT	2620
401	SerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTrr	420	QY	761	AlaSerAlaAlaSerCysAsnGlyPheLeuTrrPheValIlePhePheValAlaAlaTrr	780
1541	GGCATGTTCTCCCTTGGTGCAGGAGAGAAATCCAGCTCATCAACCAATGAGCATGG	1600	DB	2621	GGAGCGGAGCTAGCTGCAATGGCTCTCTATACTTGTTCATCTTTTCTGGCTGTGG	2680
421	HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer	440	QY	781	TrrIleLysGlyArgValValProLeuAlaThrTrrSerLeuThrGlyLeuTrrPhe	800
			DB	2681	TACATCAAGGGTGGGTAGTCCCTTGGCTACTTATTCCTTCACTGGCCTATGGTCTT	2740

QY	801	SerLeuLeuLeuLeuAlaLeuProGlnGlnAlaTyrAlaTyrAspAlaSerValHisGly	820	QY	1161	ProLeuSerThrLeuLysGlySerSerGlyGlyProValLeuCysProArgGlyHisAla	1180
DB	2741	GGCTTACTGCTCTAGCATTTGCCCAACAGCTTATGCTTATGACGATCTGTACATGGT	2800	DB	3821	CCTCTTTCCACCTTGAAGGGGCTCTCGGGGGGCGGGTCTCTGCCCAAGAGCCACGCT	3880
QY	821	GlnIleGlyAlaAlaLeuLeuValMetIleThrLeuPheThrLeuThrProGlyTyrLys	840	QY	1181	ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle	1200
DB	2801	CAGATAGGAGAGCTCTGTTGGTACTGATCCTCTTTACATCTACCCCGGGTATAG	2860	DB	3881	GTCGGGGTCTCCGGGAGCCGCTGCTCCGGGGGCGTGGCCAAAGTCCATAGATTTTATC	3940
QY	841	ThrLeuLeuSerArgPheLeuThrTrpLeuCysTyrIleuLeuThrLeuGlyGluAlaMet	860	QY	1201	ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro	1220
DB	2861	ACCCTTCTACCGCGGTTCTGCTGGTGTGCTATCTTCTGACCTTGGCGGAGCTATG	2920	DB	3941	CCCGTTGAGACACTTGACATCGTCACTCGGTCCTCCCGTCTTTAGTGACAAACAGCACCA	4000
QY	861	ValGlnGluTrpAlaProMetGlnValArgGlyGlyValArgAspGlyIleIleTrpAla	880	QY	1221	ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLys	1240
DB	2921	GTCAGAGAGTGGGACCACTATGACAGTGGCGGGTGGCGGTGATGGGATCATATGGCC	2980	DB	4001	CCTGCTGTGCCCCCAACTTATCAGGTCCGGTACTTACATGCCCGCTGGTAGGAAG	4060
QY	881	ValAlaIlePheTyrProGlyValValPheAspIleThrLysTrpLeuLeuAlaValLeu	900	QY	1241	SerThrLysValProValAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnPro	1260
DB	2981	GTGCCATATCTGCCCGGTGTGGTGTGTGACATACCAAGTGGCTCTTGGCGGTGCTT	3040	DB	4061	AGCACCAAGTCCCTGCTCGGTATGCCGTCAAGGGTACAAAGTGTAGTGTAAATCCC	4120
QY	901	GlyProAlaTyrLeuLeuLysGlyAlaLeuThrArgValProTyrPheValArgAlaHis	920	QY	1261	SerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerLysAlaHisGlyIleAsnPro	1280
DB	3041	GGGCTTCTTATCTCTAAAGGTGCTTTGACGCGTGTGCCGTACTTCTGTCAGGGCTCAC	3100	DB	4121	TCGGTGGCTGCCACCCCTGGGGTTTGGGGGTACTTGTCTCAAGGCATATGCATCAATCCC	4180
QY	921	AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyGlyArgTyrValGlnMet	940	QY	1281	AsnIleArgThrGlyValArgThrValThrThrGlyAlaProIleThrTyrSerThrTyr	1300
DB	3101	GCTCTACTAGGATGTGCACCATGGTAAAGCATCTCGCGGGGGTGGTACGTCCAGATG	3160	DB	4181	ACATTTAGGACTGGGGTCAGGACTGTGACGACCGGGGGCCCATCATCATCATAT	4240
QY	941	AlaLeuLeuAlaLeuGlyArgTyrThrGlyThrTyrIleTyrAspHisLeuThrProMet	960	QY	1301	GlyLysPheLeuAlaAspGlyGlyCysAlaGlyAlaTyrAspIleIleIleCysAsp	1320
DB	3161	GTGCTACTAGCCCTTGGCAGGTGGACTGGCATTACATCTATGACACCTCACCCCTATG	3220	DB	4241	GGCAAAATCTCTCCCGCATGGGGCTCGCGAGCGGGCTATGACATCATCATATCGCAT	4300
QY	961	SerAspTrpAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIleIlePhe	980	QY	1321	GluCysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGlnAla	1340
DB	3221	TCGGATTGGGCTGCTAATGCTCGCGGACTTGGCGGTGCGCGTGGAGCCATCATCTTC	3280	DB	4301	GAATGCCATGCCGTGACTCTACCATCTCTCGGCATCGGAACAGTCTCTCGATCAAGCA	4360
QY	981	SerProMetGluLysLysValIleValTrpGlyAlaGluThrAlaAlaCysGlyAspIle	1000	QY	1341	GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProGlySerValThr	1360
DB	3281	AGTCCGATGAGAAAAAAGTCATCGTCTGGGGAGCGGAGACAGCTGCTCGGGGATATC	3340	DB	4361	GAGACAGCCGGGGTCAGGCTNACTGTACTGGCTACGGCTACGCCCCCGGGTCAGTGACA	4420
QY	1001	LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp	1020	QY	1361	ThrProHisProAsnIleGluGluValAlaLeuGlyGlnGluGlyGluIleProPheTyr	1380
DB	3341	TTACACGGACTTCCCGTGTCCGCCGACTTGGCGGGAGTCTCTCTTGGCCAGCTGAT	3400	DB	4421	ACCCCCACCCCAACATAGAGAGGTGGCCCTCGGGCAGGAGGTGAGATTCCTTCTAT	4480
QY	1021	GlyTyrThrSerLysGlyTrpSerLeuLeuAlaProIleThrAlaTyrAlaGlnGlnThr	1040	QY	1381	GlyArgAlaIleProLeuSerTyrIleLysGlyValArgHisLeuIlePheCysHisSer	1400
DB	3401	GGCTATACCTCCAGGGGTGGAGTCTTCTCGCCCCCATCTGCTTATGCCCCAGCAGACA	3460	DB	4481	GGAGGGGCGATTCCTCTGTATACATCAAGGGAGGAAGACACTTGATCTTCTGCCACTCA	4540
QY	1041	ArgGlyLeuLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGluGlnAla	1060	QY	1401	LysLysLysCysAspGluLeuAlaAlaLeuArgGlyMetGlyLeuAsnSerValAla	1420
DB	3461	CGCGGCTTTTGGGACCATAGTGGTGGATGACGGGGCGGCAAGACAGAACAGGCC	3520	DB	4541	AGAAAAAGTGTGACGAGCTCGCGCGGCCCTTTCGGGGTATGGCTTGAACGAGTGGCA	4600
QY	1061	GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly	1080	QY	1421	TyrTyrArgGlyLeuAspValSerValIleProThrGlnGlyAspValValValAla	1440
DB	3521	GGGAGATTTAGGTCCTGTCCACGGTCACTCAGTCTCTTCTCGGAAACAAACCTCGGG	3580	DB	4601	TACTACAGAGGGTGGACGCTCTCCGTAATACCAACTCAGGAGACGTAGTGGTCTGCC	4660
QY	1081	ValLeuTrpThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro	1100	QY	1441	ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnVal	1460
DB	3581	GTCCTATGAGATGTCTACCTAGGCTGGCAACAGACTCTAGCGGCTCACGGGGTCCG	3640	DB	4661	ACCGAGCCCTCATGACGGGGTTTCTGGAGACTTTGACTCCGCTGATCGACTGCCACGTA	4720
QY	1101	ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTrpProSerProProGly	1120	QY	1461	AlaValThrGlnValValAspPheSerLeuAspProThrPheThrIleThrThrGlnIle	1480
DB	3641	GTCACAGAGTGTACTCAGTGTCTGAGGGGAGCTTAGTGGGGTGGCCAGCCCCCGGG	3700	DB	4721	CGGGTCACTCAAGTTGTAGACTTCAGCTTGGACCCCACTTACCATACCAACACAGACT	4780
QY	1121	ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArgAsn	1140	QY	1481	ValProGlnAspAlaValSerArgSerGlnArgGlyArgThrGlyArgGlyArgLeu	1500
DB	3701	ACCAAACTTTGGAGCGGTGACCGTGTGGAGCGGTCCAGCTATACCTTGGTCCAGGAAC	3760	DB	4781	GTCCCTCAAGACGCTGCTCTCACGTAGCCAGCGCGGGCGCCACGGGACGGGGAAGACTG	4840
QY	1141	AlaAspValIleProAlaArgArgGlyAspLysArgGlyAlaLeuLeuSerProArg	1160	QY	1501	GlyIleTyrArgTyrValSerThrGlyGluArgAlaSerGlyMetPheAspSerValVal	1520
DB	3761	GCTGATGTCTCCCGGCTCGAAGACGCGGGGACCAAGGAGGAGCGCTACTCTCCCGGAGA	3820	DB	4841	GGTATTATAGGTATGTTTCCACTGGTGGAGCGGCTTCAGGAATGTTTGACAGTGTAGTG	4900
QY				QY	1521	LeuCysGluCysTyrAspAlaGlyAlaAlaTrpTyrGluLeuThrProSerGluThrThr	1540

Db	4901	CTCTGCGAGTGCTACGATGCGGGCGCGCATGCTGATGAGCTCAACAGCGGAGACCAACC	4960	Db	5981	GAGGATGTTGTCAACCTGCTGCTCGGAATTTCTGTCTCCGGGTGCTGCTGGTGGGAGTC	6040
Qy	1541	ValArgLeuArgAlaTyrPheAsnThrProGlyLeuProValCysGlnAspHisLeuGlu	1560	Qy	1901	IleCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTyrMet	1920
Db	4961	GTGAGGCTCAGAGCATATTTCAACACACACCTGGTTTGGCTGTGTGCGCAAGACCATTTGAG	5020	Db	6041	ATCTGCGCGCCCATCTCTGCGCGACACAGTGGGACCGGGGGAGGCGCTGTCCAAATGGATG	6100
Qy	1561	PheTrpGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThr	1580	Qy	1921	AsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaProThrHisIleValThr	1940
Db	5021	TTTTTGGGAGGAGTTTTCACGGGCTCACACATAGATGCCCACTTCTTTTCCCAACA	5080	Db	6101	AATAGGCTCATTTGCTTCTCCAGAGGAACACAGTGCGCCCCACCCACTACGTGACG	6160
Qy	1581	LysGlnSerGlyGluAsnPheAlaTyrLeuThrAlaTyrGlnAlaThrValCysAlaArg	1600	Qy	1941	GluSerAspAlaSerGlnArgValThrGlnLeuLeuGlySerLeuThrIleThrSerLeu	1960
Db	5081	AAGCAATCGGGGGAATAATTCGCATACTTAACAGCCTTACCAGGCTACAGTGTGCGGTAGG	5140	Db	6161	GAGTCCGATCGCTGCGAGCGTGTACCCAACTACTTTGGCTCTCTTACCAATAACAGCGCTG	6220
Qy	1601	AlaIleAlaProProSerTrpAspValMetTrpLysCysLeuThrArgLeuLysPro	1620	Qy	1961	LeuArgArgLeuHisAsnTrpIleThrGluAspCysProIleProCysGlyGlySerTrp	1980
Db	5141	GCCAAAGCCCCCCCCCTCTGGAGCGTCATGTGGAAGTGTTTGACTCGACTCAAGCCC	5200	Db	6221	CTCAGAGACTCCCAACTGGATTACTGAAGACTGCCCCATCCCATGCGAGCGGCTGTGG	6280
Qy	1621	ThrLeuValGlyProThrProLeuLeuTyrArgLeuGlySerValAsnGluValThr	1640	Qy	1981	LeuArgAspValTrpAspTrpValCysThrIleLeuThrAspPheLysAsnTrpLeuThr	2000
Db	5201	ACACTGCTGGGCCCCACACCTCTCTGTACCGCTTGGGCTCTGTACCAACGAGGTCAAC	5260	Db	6281	CTCCGCGATGTGGGATTTGGGTTTGCCATCTTAACAGACTTTAAAAAAGTGGTGACC	6340
Qy	1641	LeuThrHisProValThrLysTyrIleAlaThrCysMetGlnAlaAspLeuGluValMet	1660	Qy	2001	SerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTyrLys	2020
Db	5261	CTCAGGCATCTGTGACCAATAATCATCCACCTGCATGCAAGCCGACCTTGAGGTCAATG	5320	Db	6341	TCCAAATTTGTTCCAAAGATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	6400
Qy	1661	ThrSerThrTrpValLeuAlaGlyGlyValLeuAlaAlaValAlaTyrCysLeuAla	1680	Qy	2021	GlyValTrpAlaGlyThrGlyIleMetThrThrArgCysProCysGlyAlaAsnIleSer	2040
Db	5321	ACCAGCAGTGGGTCTTAGCTGGGGGGCTCTTGGCGCGCTCGCCGCTACTGCCCTGGCG	5380	Db	6401	GGCGTGTGGCTGGCACTGCTGATCATCACACAGGTGCTCTTGGCGGCGCAATATCTCT	6460
Qy	1681	ThrGlyCysValCysIleIleGlyArgLeuHisIleAsnGlnArgAlaValAlaPro	1700	Qy	2041	GlyAsnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTrp	2060
Db	5381	ACCGGTGTGTGTCATCATCGCGCTTGCAGTTTAAACAGCGAGCGCTGTGTGACCCG	5440	Db	6461	GGCAATGTCGCTCGCTCCATGAGAAATACGGGGCCCAAAACCTGCATGAATATCTGG	6520
Qy	1701	AspLysGluValLeuTyrGluAlaPheAspGluMetGluGluCysAlaSerArgAlaAla	1720	Qy	2061	GlnGlyThrPheProIleAsnCysTyrThrGluGlyGlnCysValProLysProAlaPro	2080
Db	5441	GACAAGGAGTCTCTATAGGCTTTTGTATGATGAGTGGAGAAATGTGCTCTAGAGCGCT	5500	Db	6521	CAGGGAGCTTTTCCCATCAATTTTACACGGAGGCGAGTGCCTGCGAAACCGCACCA	6580
Qy	1721	LeuIleGluGlyGlnArgIleAlaGluMetLeuLysSerLysIleGlnIleLeuLeu	1740	Qy	2081	AsnPheLysValAlaIleTrpArgValAlaAlaSerGlyTyrAlaGluValThrGlnHis	2100
Db	5501	CTCATTAAGAGGGGCGAGGATAGCGAGATGCTGAAGTCCAAAGATCCAAAGCTTATTG	5560	Db	6581	AACTTTAAGATCCCATCTGGAGGTGGCGCTCAGAGTACCGGAGGTGACGAGCAC	6640
Qy	1741	GlnGlnAlaSerLysGlnAlaGlnAspIleGlnProThrValGlnAlaSerTrpProLys	1760	Qy	2101	GlySerTyrHisIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeu	2120
Db	5561	CAGCAAGTTTCCAAACAGCTCAGACATACACCCGCTGTGAGGCTTCTTGGCCCAAG	5620	Db	6641	GGGTATACACCTTACATAACAGGACTTACCACTGATAACTTGAAGTTCTCTGCCAATA	6700
Qy	1761	ValGluGlnPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeuAla	1780	Qy	2121	ProSerProGluPhePheSerTrpValAspGlyValGlnIleHisArgPheAlaProThr	2140
Db	5621	GTAGAGCAATTCTGGGCCAAACACATGTGGAACCTTCATCAGCGGCATTCAATACCTGCA	5680	Db	6701	CCTTCTCCAGAGTTCTTTCTCTGGGTGACGGAGTGCAGATCCATAGGTTTGGCCCCATA	6760
Qy	1781	GlyLeuSerThrLeuProGlyAsnProAlaValAlaSerMetMetAlaPheSerAlaAla	1800	Qy	2141	ProLysProPhePheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal	2160
Db	5681	GSACTATCAACACTGCCAGGAAACCTCTGTGTAGCTTCCATGATGGCATTCAGTGGCGCC	5740	Db	6761	CCGAACCGTTTTTTCGGGATGAGGTCTGCTTCTGGTGGGCTTAAATCATTTGTGCTC	6820
Qy	1801	LeuThrSerProLeuSerThrSerThrIleLeuLeuAsnIleLeuGlyGlyTrpLeu	1820	Qy	2161	GlySerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThr	2180
Db	5741	CTCACCAGTCCGTGTCACTAGCACCACTATCTCTTCAACATTTTGGGGGGCTGGCTA	5800	Db	6821	GGGTCTCAGCTCCCTTGGCATCTCTGAACCTGCACAGACGATTTGACGTCTCATTAACA	6880
Qy	1821	AlaSerGlnIleAlaProProAlaGlyAlaThrGlyPheValValSerGlyLeuValGly	1840	Qy	2181	AspProSerHisIleThrAlaGluThrAlaAlaArgArgLeuAlaArgGlySerProPro	2200
Db	5801	GCATCCCAATTTGCGCTCCCGGGGGGTACCGGCTTCTGCTGAGTGGTGGTGGG	5860	Db	6881	GACCCATCCCATATACGCGGAGACTGACAGCGCGGCTTTGGCACGGGGTCAACCCCG	6940
Qy	1841	AlaAlaValGlySerIleGlyLeuGlyLysValLeuValAspIleLeuAlaGlyTyrGly	1860	Qy	2201	SerGluAlaSerSerAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThr	2220
Db	5861	GCTGCCGTAGCAGCATAGGCTTGGGTAAAGTGTGTTGGTGGACATCTCTGGCAGGATGGT	5920	Db	6941	TCCGAGGCAAGCTCTTCAGCGAGCAGTATCGCACCATCTGCTGGAGGCCACTGACCC	7000
Qy	1861	AlaGlyIleSerGlyAlaLeuValAlaPheLysIleMetSerGlyGluLysProSerMet	1880	Qy	2221	ThrHisGlyLysAlaTyrAspValAspMetValAspAlaAsnLeuPheMetGlyGlyAsp	2240
Db	5921	GCGGGCATTTTGGGGGCTCTGCTGCGCATTAAGATCATGTGCGGAGAGCCCTCCATG	5980	Db	7001	ACCCAGCGGACGCTTATGATGGACATGGTGGATGCCCACTGTTTCATGGGGGGCAT	7060
Qy	1881	GluAspValValAsnLeuLeuProGlyIleLeuSerProGlyAlaLeuValValGlyVal	1900	Qy	2241	ValThrArgIleGluSerGlySerLysValValValLeuAspSerLeuAspProMetVal	2260
Db				Db	7061	GTGACCGGATAGAGTCTGAGTCCAAAGTGGTCTGACTCTCTGAGCTCTCTCGACCAATGGTC	7120

2261	Qy	GluGluArgSerAspLeuGluProSerIleProSerGluTyrMetLeuProLysIysArg	2280
7121	Ddb	GAAGAAAGGAGCGACTTGAGCCTTCGATACCATCGGAATAATATGCTCCCAAGAAGAGA	7180
2281	Qy	PheProProAlaLeuProAlaTrpAlaArgProAspTyrAsnProProLeuValGluSer	2300
7181	Ddb	TTCCCAAGCAGCCTTACCGCTTTGGGCACGGCTGATTACAAGCCACCGCTTGTGAATCG	7240
2301	Qy	TrpLysArgProAspTyrClnProAlaThrValAlaGlyCysAlaLeuProProArg	2320
7241	Ddb	TGGAAGAGGCCAGATTACCAACCGGCCACTGTGTGGGGCTGCGCTCTCCCCCCCTTAAG	7300
2321	Qy	LysThrProThrProProArgArgArgThrValGlyLeuSerGluAspSerIle	2340
7301	Ddb	AAAAACCCAGCGCTCCCCCAAGGAGACGCCCGACAGTGGGTCTGAGTGAGAGCTCCATA	7360
2341	Qy	GlyAspAlaLeuGlnGlnLeuAlaIleTysSerPheGlyGlnProProProSerGlyAsp	2360
7361	Ddb	GCAGATGCCCTACAACAGCTGGCCATCAAGTCTCTTTGGCCAGCCGCCCCCAAGCGCGAT	7420
2361	Qy	SerGlyLeuSerThrGlyAlaGlyAlaAlaAspSerGlySerGlnThrProProAspGlu	2380
7421	Ddb	TCAGGCCTTTCCACGGGGCGGACGCGAGCGGATTCGGCAGTCGGACGCCCCCGATGAG	7480
2381	Qy	LeuAlaLeuSerGluThrGlySerIleSerSerMetProProLeuGluGlyGluLeuGly	2400
7481	Ddb	TTGGCCCTTTCCGAGACAGGTTCATCTCCTCATGCCCTCTCTGAGGGGAGCGCTGGA	7540
2401	Qy	AspProAspLeuGluProGluGlnValGluProGlnProProGlnGlyValAla	2420
7541	Ddb	GATCAGACCTTGAGGCTTCAGCAGGTAGAGCTTCAACCTCCCCCAGGGGGGTGTA	7600
2421	Qy	AlaProGlySerAspSerGlySerTrpSerThrCysSerGluGluAspAspSerValVal	2440
7601	Ddb	ACCCCGCGCTCAGGCTCGGGGCTTGGTCTACTTGTCTCGAGGAGACGACTCCGTCGTG	7660
2441	Qy	CysCysSerMetSerTyrSerTrpThrGlyAlaLeuIleThrProCysSerProGluGlu	2460
7661	Ddb	TGCTGCTCCATGTCATCTCTCGACCGGGGCTCTAATACTCTCTGTAGCCCCGAGAG	7720
2461	Qy	GluLysLeuProIleAsnProLeuSerAsnSerLeuLeuArgTyrHisAsnLysValTyr	2480
7721	Ddb	GAAAGTTGCCAATTAACCCCTTGACCAACTCCCTGTTGGGATATCAACAAGGTGTAC	7780
2481	Qy	CysThrThrTrpLysSerAlaSerLeuArgAlaLysLysValThrPheAspArgMetGln	2500
7781	Ddb	TGTACCACATCAAGAGCGCCTCATTAAGGGCTAAAGAGGTAACTTTTGATAGGATGCA	7840
2501	Qy	ValLeuAspSerTyrTrpAspSerValLeuLysAspIleLysLeuAlaLysLysVal	2520
7841	Ddb	GGCTTCGAGCGCTCATTTAGCTCAGCTTTGAAGGACATTAAAGCTACGGCCCTCAAGGTC	7900
2521	Qy	ThrAlaArgLeuLeuThrMetGluGluAlaCysGlnLeuThrProProHisSerAlaArg	2540
7901	Ddb	ACCGAAGGCTTCTCACTTTAGAGGAGGCTGCGAGTTAACTCCACCCCACTCTGCAAGA	7960
2541	Qy	SerLysTyrGlyPheGlyAlaLysGluValArgSerLeuSerGlyArgAlaValAsnHis	2560
7961	Ddb	TCCAAGTATGGGTTTGGGCTTAAGGAGGTCGCGAGCTTGTCCGGGAGAGCCGTTAACCCAC	8020
2561	Qy	IleLysSerValTrpLysAspLeuLeuLysAspSerGluThrProIleProThrThrIle	2580
8021	Ddb	ATCAAGTCGCTGGGAAGACCTCTCGAAGACACACAACCAATCTCTCAACCACTC	8080
2581	Qy	MetAlaLysAsnGluValPheCysValAspProThrLysGlyGlyLysLysAlaLysArg	2600
8081	Ddb	ATGGCCAAAATGAGGTGTTCTCGTGGACCCCAAGGGGGGTAAAGACGAGCTCGC	8140
2601	Qy	LeuIleValTyrProAspLeuGlyValArgValCysGluLysMetAlaLeuTyrAspIle	2620
8141	Ddb	CTTATCGTTTACCCTTGACCTCGCGCTCAGGGTCTGCGAGAAATGCCCCCTTTATGATATC	8200

2621	ThrGlnLysLeuProGlnAlaValMetGlyAlaSerTyrGlyPheGlnTyrSerProAla	2640
2622		2641
2623		2642
2624		2643
2625		2644
2626		2645
2627		2646
2628		2647
2629		2648
2630		2649
2631		2650
2632		2651
2633		2652
2634		2653
2635		2654
2636		2655
2637		2656
2638		2657
2639		2658
2640		2659
2641		2660
2642		2661
2643		2662
2644		2663
2645		2664
2646		2665
2647		2666
2648		2667
2649		2668
2650		2669
2651		2670
2652		2671
2653		2672
2654		2673
2655		2674
2656		2675
2657		2676
2658		2677
2659		2678
2660		2679
2661		2680
2662		2681
2663		2682
2664		2683
2665		2684
2666		2685
2667		2686
2668		2687
2669		2688
2670		2689
2671		2690
2672		2691
2673		2692
2674		2693
2675		2694
2676		2695
2677		2696
2678		2697
2679		2698
2680		2699
2681		2700
2682		2701
2683		2702
2684		2703
2685		2704
2686		2705
2687		2706
2688		2707
2689		2708
2690		2709
2691		2710
2692		2711
2693		2712
2694		2713
2695		2714
2696		2715
2697		2716
2698		2717
2699		2718
2700		2719
2701		2720
2702		2721
2703		2722
2704		2723
2705		2724
2706		2725
2707		2726
2708		2727
2709		2728
2710		2729
2711		2730
2712		2731
2713		2732
2714		2733
2715		2734
2716		2735
2717		2736
2718		2737
2719		2738
2720		2739
2721		2740
2722		2741
2723		2742
2724		2743
2725		2744
2726		2745
2727		2746
2728		2747
2729		2748
2730		2749
2731		2750
2732		2751
2733		2752
2734		2753
2735		2754
2736		2755
2737		2756
2738		2757
2739		2758
2740		2759
2741		2760
2742		2761
2743		2762
2744		2763
2745		2764
2746		2765
2747		2766
2748		2767
2749		2768
2750		2769
2751		2770
2752		2771
2753		2772
2754		2773
2755		2774
2756		2775
2757		2776
2758		2777
2759		2778
2760		2779

```

Db      9281  TTCCGGAAGCGCGCTCTCGGATTTATCCAGCTGTTCACTGTCGCGCGCGCGGGC 9340
Qy      3001  AspileYrHisSerValSerArgAlaArgProArgLeuLeuLeuPheGlyLeuLeu 3020
Db      9341  GACATTATACAGCGTGTGCGTGGCCGACCGCTTATCTCTTGGCCCTACTCTTA 9400
Qy      3021  LeuPheValGlyValGlyLeuPheLeuLeuProAlaArg 3033
Db      9401  CTTTGTAGGGTAGGCCCTTTCTACTCTCCCGCTCGG 9439

RESULT 6
E07361
LOCUS      E07361          9589 bp    RNA          linear    PAT 29-SEP-1997
DEFINITION GRNA of Hepatitis non-A non-B virus, HCV-J6.
ACCESSION  E07361
VERSION     E07361.1 GI:5708558
KEYWORDS   JP 1994121689-A/1.
SOURCE      unidentified
ORGANISM    unclassified.
REFERENCE   1 (bases 1 to 9589)
AUTHORS     Okamoto,H. and Nakamura,T.
TITLE       NON-A NON-B HEPATITIC VIRUS GENE, POLYNUCLEOTIDE, POLYPEPTIDE,
            ANTIGEN AND ANTIBODY DETECTION SYSTEM
JOURNAL     Patent: JP 1994121689-A 1 06-MAY-1994;
            NAKAMURA TETSUO
COMMENT     OS Hepatitis non-A non-B virus
            PN JP 1994121689-A/1
            PD 06-MAY-1994
            PF 09-AUG-1991 JP 1991287402
            PI OKAMOTO HIROAKI, NAKAMURA TETSUO
            PC C12N15/51,A61K39/29,A61K39/395,C07K13/00,C12P21/02,
            PC C12P21/08,
            PC G01N33/53,G01N33/576,G01N33/577//A61B10/00;
            CC strandedness: Single;
            CC topology: Linear;
            CC hypothetical: No;
            CC anti-sense: No;
            FH Key
            FT source
            FT 1..9589
            FT /organism='Hepatitis non-A non-B virus', FT
            FT /strains='HCV-J6'.
            FT /Location/Qualifiers
FEATURES
source      1..9589
            /organism='unidentified'
            /mol_type='genomic RNA'
            /db_xref='taxon:32644'
BASE COUNT  1968 a 2820 c 2635 g 2166 t
ORIGIN
Alignment Scores:
Pred. No.:      0
Score:          15824.00
Percent Similarity: 98.78%
Best Local Similarity: 97.79%
Query Match:    98.16%
DB:             6
            Length:      9589
            Matches:    2966
            Conservative: 30
            Mismatches:  37
            Indels:      0
            Gaps:        0

US-09-980-559-2 (1-3033) x E07361 (1-9589)
Qy      1  MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgProGln 20
Db      341  ATGAGCAAAATCTTAACCTCAAGAAACCAAGAAACCAACCAACCAACCAACCAAC 400
Qy      21  AspValLysPheProGlyGlyGlnLeuValGlyGlyValTyrLeuLeuProArgArg 40
Db      401  GACGTTAAGTTTCGGCGCGCGCCAGATCGTTGGCGAGTATATCTTGTTCGCCGAGG 460
Qy      41  GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly 60
Db      461  GGCCTCCAGGTTGGTGTGTCGCGGCAAGGAAGACTTCGAGCGGTTCGCCAGCCAGTGG 520
Qy      61  ArgArgGlnProLysAspArgArgSerThrGlyLysSerTrpGlyLysProGly 80
Db      521  AGCGCGCGAGCCCATCCCTAAGGATCGCGCTCCACTGGCAATCTCTGGGGAACACGGA 580
Qy      81  TyrProTrpProLeuLeuTyrGlyAsnGluGlyLeuGlyTrpAlaGlyTrpLeuLeuSerPro 100
Db      581  TACCCCTGGCCCCATACCGGAATAGAGGACTCGGCTGGGCGAGGATGGCTCTCTGTCCCC 640
Qy      101  ArgGlySerArgProSerTrpGlyProAsnAspProArgHisArgSerArgAsnValGly 120
Db      641  CGAGGTTCCGCTCCCTTTGGGGCCCCAATAGCCCCCGCATAGGTCCTCCCAACAGTGGGT 700
Qy      121  LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProValVal 140
Db      701  AAGGTCATCGATACCTTAACGTGGCGCTTTGCCGACCTCATGGGGTACATCCCTGTGCTA 760
Qy      141  GlyAlaProLeuGlyGlyValAlaArgAlaLeuAlaHisGlyValArgValLeuGluAsp 160
Db      761  GCGCGCCCGCTCGCGCGCGCTCGCAGAGCTCTCGCGCATGGCGTGAGAGTCTCTGGAGGAC 820
Qy      161  GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla 180
Db      821  GGGGTTAAATTTGCAACAGGGAACCTTACCCTGGTGTCTCTCTTCTATCTCTTGTGGGCC 880
Qy      181  LeuLeuSerCysIleThrThrProValSerAlaAlaGluValLysAsnIleSerThrGly 200
Db      881  CTGCTGTCTCGCATCACACCCCGGTCTCGCTGCGCGAAGTGAAGAACATCAGTACCGGC 940
Qy      201  TyrMetValThrAsnAspCysThrAsnAspSerIleThrTrpGlnLeuGlnAlaVal 220
Db      941  TACATGGTCACCAACGACTGCACCAATGATAGCATTTACCTGGCACTCCAGGCTGTCTGC 1000
Qy      221  LeuHisValProGlyCysValProCysGluLeuValGlyAsnAlaSerGlnCysTrpIle 240
Db      1001  CTCACGTCCTCCGGGTGCGTCCGTCGAGAAAGTGGGGAATACATCTCGGTCTGTGATA 1060
Qy      241  ProValSerProAsnValAlaValGlnArgProGlyAlaLeuThrGlnGlyLeuArgThr 260
Db      1061  CCGGTCTCACCGAATGTCGCGTGCAGCAGCCCGCGCTTCACGCGAGGCTTACCGAGC 1120
Qy      261  HisIleAspMetValValMetSerAlaThrLeuCysSerAlaLeuTyrValGlyAspLeu 280
Db      1121  CACATTGACATGGTGTGATGTCGCGCAGCTCTGCTCGCTCTTTACGTGGGGGACCTC 1180
Qy      281  CysGlyGlyValMetLeuAlaGlnMetPheIleValSerProGlnHisHisTrpPhe 300
Db      1181  TCGCGTGGGGTGATGCTTGACGCCAGATGTTTCATTGTCTCGCCACAGCACCACTGGTTT 1240
Qy      301  ValGlnAspCysAsnCysSerIleTyrProGlyThrIleThrGlyHisArgMetAlaTrp 320
Db      1241  GTGCAGACTGCATTTGCTCCATCTACCTTGGTACCATCTACGGACACCGCATGGGGTGG 1300
Qy      321  AspMetMetMetAsnTrpSerProThrAlaThrMetIleLeuAlaTyrAlaMetArgVal 340
Db      1301  GACATGATGATGAATGGTTCGCGCGCTACCATGATCTTGGCGTACGCGATGCGGCTC 1360
Qy      341  ProGluValIleIleAspIleIleSerGlyAlaHisTrpGlyValMetPheGlyLeuAla 360
Db      1361  CCGGAGGTTCATATGACATCATTTGGCGGGGTTCATTGGGGCGTCAATGTTGGCTTAGCC 1420
Qy      361  TyrPheSerMetGlnGlyAlaTrpAlaLysValValIleLeuLeuLeuAlaAlaGly 380
Db      1421  TACTTCTCTATGCGAGGAGCGTGGGCAAAAGTCTGTGTCTATCTTTTGTGGCGCGGG 1480
Qy      381  ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr 400
Db      1481  GTGGACGCGCAAAACCCATACCGTGGGGGTTCCTACCGCGCAATAACGCCAGGACCTCACC 1540
Qy      401  SerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTrp 420
Db      1541  GGCATGTTCTCCTTGGTCCAGGCAAAATCCAGCTCATCAACCAATGCGAGTGG 1600

```

2681	TACATCAGGGTCGGGTAGTCCCTTGGGTACTTATTTCCTCCACTGGCCTATGGTCTCTT	2740
801	SerLeuLeuLeuAlaLeuProGlnAlaTyrAlaTyrAspAlaSerValHisGly	820
2741	GGCTACTGCTCTAGCATTGCCCAACAGGCTTATGCTTATGACGCATCTGTACATGGT	2800
821	GlnIleGlyAlaAlaLeuValMetIleThrLeuPheThrLeuThrProGlyTyrLys	840
2801	CAGATAGGAGCAGCTCTGTGTGTACTCATCTCTTTTACACTCACCCCGGGTATAAG	2860
841	ThrLeuLeuSerArgPheLeuTrpTrpLeuCyfThrLeuThrLeuGlyGlulAlaMet	860
2861	ACCTCTTCTCAGCCGGTCTCTGTGGTGTGTGCTATCTCTGACCTTGGCGAAGCTATG	2920
861	ValGlnGluTrpAlaProMetGlnValArgGlyGlyArgAspGlyIleIleTrpAla	880
2921	GTCCAGAGTGGCACCACTTATGCAGGTGGCGGTGGCGTGTATGGGATCATATGGCC	2980
881	ValAlaIlePheTyrProGlyValValPheAspIleThrLysTrpLeuLeuAlaValLeu	900
2981	GTGCCATATCTGCCGGGTGTGTGTGTTGACATAACCAAGTGGCTCTTGGCGGTGCTT	3040
901	GlyProAlaTyrLeuLeuLysGlyAlaLeuThrArgValProTyrPheValArgAlaHis	920
3041	GGGCCCTGCTTATCTCTAAAGGTGCTTGTGACGGGTGCGGTACTTCTGCAGGGCTCAC	3100
921	AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyGlyArgTyrValGlnMet	940
3101	GCTCTACTAAGGATGTGCACCATGTAAGGATCTCCGGGGGGTAGTACGTCCAGATG	3160
941	AlaLeuLeuAlaLeuGlyArgTrpThrGlyThrTyrIleTyrAspHisLeuThrProMet	960
3161	GTGCTACTAGCCCTTGGCAGGTGGACTGGCACTTACATCTATGACCACTTCACCCCTATG	3220
961	SerAspTrpAlaAlaSerClyLeuArgAspLeuAlaValAlaValGluProIleIlePhe	980
3221	TCGGATTGGGCTGCTAATGGCTCGCGGACTTGGCGGTGGCGGTGGAGCCCTATCATCTTC	3280
981	SerProMetGluLysLysValIleValTrpGlyAlaGluThrAlaAlaCysGlyAspIle	1000
3281	AGTCCCATGGAGAAAMAAGTCACTGCTGGGAGCGGAGACAGCTGCTGGGGGATATC	3340
1001	LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp	1020
3341	TTACAGGACTTCCCGTGTCCGCCGACTTGGCCGGGAGGTCTCTCTTGGCCAGCTGAT	3400
1021	GlyTyrThrSerLysGlyTrpSerLeuLeuAlaProIleThrAlaTyrAlaGlnGlnThr	1040
3401	GGCTATACCTCCAAGGGGTGGAGTCTCTCGCCCCCATCACTGCTATATGCCAGCAGACA	3460
1041	ArgGlyLeuLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGluGlnAla	1060
3461	CGCGCCCTTTGGGCACCATAGTGTGTAGCATGACGGGGCGCAGACAGAACAGGCC	3520
1061	GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly	1080
3521	GGGAGATTCCAGTCTGTGCCAGTCACTCAGTCTCTCTCTCGGAACAACCATCTCGGG	3580
1081	ValLeuTrpThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro	1100
3581	GTCTTATGGACTGCTCATACCATGGAGCTGGCAACAAGACTCTAGCCCGGCTCACCGGGTCCG	3640
1101	ValThrLysMetTyrSerSerAlaGluGlyAspLeuValGlyTyrProSerProProGly	1120
3641	GTCAACAGATGTACTCCAGTCTGAGGGGGACTTAGTGGGGTGGCCAGCCCCCGGG	3700
1121	ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArgAsn	1140
3701	ACCAATCTTTGGAGCCGTGCAGTGTGAGCGGTGCACTTATACCTGGTCAAGCAAC	3760
1141	AlaAspValIleProAlaArgArgGlyAspLysArgGlyAlaLeuLeuSerProArg	1160

Db	3761	GCTGATGTCTATCCCGCTCGAAGACGCGGGCAACAGCGAGCGGTACTCTCCCCGAGA	3820	QY	1521	LeuCysGluCysTyrAspAlaGlyAlaAlaTptTyrGluLeuThrProSerGluThrThr	1540
QY	1161	ProLeuSerThrLeuLysGlySerSerGlyGlyProValLeuCysProArgGlyHisAla	1180	Db	4901	CTCTGGAGTGTACGATCGAGGGCGCGATGGTATGAGTCACACCGAGGAGACCACC	4960
Db	3821	CTCTTTTCCACCTTGAGGGGTCTCTGGGGGGCCCGGTGCTCTGCCCGAGGCGCAGCT	3880	QY	1541	ValArgLeuArgAlaTyrPheAsnThrProGlyLeuProValCysGlnAspHisLeuGlu	1560
QY	1181	ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle	1200	Db	4961	GTCAGGCTCAGAGCATATTTCAACACACCTGGTTTGCCTGTGTGCCAAGACCATCTTGAG	5020
Db	3881	GTCGGGGTCTTCGGGGCAGCGGTGTCTCCGGGGCGGTGGCCAAAGTCCATAGATTTATC	3940	QY	1561	PheTrpGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThr	1580
QY	1201	ProValGlnThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro	1220	Db	5021	TTTTTGGAGGCGAGTTTTTACCAGCCTCACACATAGATGCCACCTTCCTTTTCCCAACA	5080
Db	3941	CCCGTTGAGACACTTGACATCGTCACCTCGGTCCCCCACCCTTTAGTGACACACACCA	4000	QY	1581	LysGlnSerGlyGluAsnPheAlaTyrLeuThrAlaTyrGlnAlaThrValCysAlaArg	1600
QY	1221	ProAlaValProGlnThrTyrGlnValGlyLeuHisAlaProThrGlySerGlyLys	1240	Db	5081	NAGCAATCGGGGMAAATTTCCGACATTTAACAGCCTACCAGGCTACAGTGTGGCTAGG	5140
Db	4001	CTGTGCTGTGCCCAACTTATCAGGTCCGGTACTTACATGCCCGCGACTGGTAGTGGAAAG	4060	QY	1601	AlaLysAlaProProProSerTrpAspValMetTrpLysCysLeuThrArgLeuLysPro	1620
QY	1241	SerThrLysValProValAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnPro	1260	Db	5141	GCCAAAGCCCCCCCCCGTCTCTGGGACGTGATGTGAAGTGTGTGACTCGACTCAAGCCC	5200
Db	4061	AGCACCAAGTCCCTGTCCGCTATGCCGTTCAGGGGTACAAAGTGTAGTGTAAATCCC	4120	QY	1621	ThrLeuValGlyProThrProLeuLeuTyrArgLeuGlySerValThrAsnGluValThr	1640
QY	1261	SerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerLysAlaHisGlyIleAsnPro	1280	Db	5201	ACACTCGTGGGCCCCACACTCTCTGTGTACCGCTTGGGCTCTGTATACCACAGAGGTACC	5260
Db	4121	TCGGTGGCTGCCACCTCGGGGTTTGGGGCGGTACTTGTCCAAGGCACATGGCATCAATCCC	4180	QY	1641	LeuThrHisProValThrLysTyrIleAlaThrCysMetGlnAlaAspLeuGluValMet	1660
QY	1281	AsnIleArgThrGlyValArgThrValThrThrGlyAlaProIleThrTyrSerThrTyr	1300	Db	5261	CTCAGCATCTCTGTGAGGAATACATCGCCACCTGCTGATGCAAGCCGACCTTGAGGTCTATG	5320
Db	4181	ACATTTAGGACTGGGTGAGGACTGTGACGCCGGGGCGCCCATCAGTACTCCACATAT	4240	QY	1661	ThrSerThrTrpValLeuAlaGlyGlyValLeuAlaAlaValAlaAlaTyrCysLeuAla	1680
QY	1301	GlyLysPheLeuAlaAspGlyGlyCysAlaGlyGlyAlaTyrAspIleIleCysAsp	1320	Db	5321	ACCAGCAGCTGGGTCTTAGCTGGGGGGTCTTGGCGCGCTGCGCGCTACTGCTTGGCGG	5380
Db	4241	GGCAAAATTCCTGCCCGATGGGGCTGCGCAGGCGCGCTTATGACATCATATGCGAT	4300	QY	1681	ThrGlyCysValCysIleIleGlyArgLeuHisIleAsnGlnArgAlaValAlaPro	1700
QY	1321	GluCysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGlnAla	1340	Db	5381	ACCGGTGTGTGTGATCATCGCCGCTTGGCAGCTTAACAGCAGCGCGCTGCTGTGACCGG	5440
Db	4301	GAATGCCATGCCGTGGACTTACACCATTTCTCGGCATCGGAACAGTCTCTCGATCAAGCA	4360	QY	1701	AspLysGluValLeuTyrGluAlaPheAspGluMetGluGluCysAlaSerArgAlaAla	1720
QY	1341	GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProProGlySerValThr	1360	Db	5441	GACAGGAGGTCTCTATGAGGCTTTTGATGAGATGAGAGGATGTGCTCTAGAGCGGCT	5500
Db	4361	GAGACAGCCGGGTGAGGCTTAACCTGTACTGGCTACGGCTACGCCCGCGCGTCAAGTGA	4420	QY	1721	LeuIleGluGlyGlnArgIleAlaGluMetLeuLysSerLysIleGlnGlyLeuLeu	1740
QY	1361	ThrProHisProAsnIleGluGluValAlaLeuGlyGlnGlyGluIleProPheTyr	1380	Db	5501	CTCATTTGAAGAGGGGCGAGCGGATAGCCGAGATGCTGAAGTCCCAAGATCCAAAGCTTATTG	5560
Db	4421	ACCCCCCCCCCAACATAGAGGAGTGGCCCTCGGCGAGGAGGTGAGATCCCTCTTAT	4480	QY	1741	GlnGlnAlaSerLysGlnAlaGlnAspIleGlnProThrValGlnAlaSerTrpProLys	1760
QY	1381	GlyArgAlaIleProLeuSerTyrIleLysGlyArgHisLeuIlePheCysHisSer	1400	Db	5561	CAGCAAGCTTCCAAACAAGCTCAAGACATACAAACCCGCTGTGAGGCTTCTTGGCCCAAG	5620
Db	4481	GGGAGGGCGATTCCCTGTATATCATATCAAGGGAGGAAGACATTTGATCTTCTGCCACTCA	4540	QY	1761	ValGluGlnPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeuAla	1780
QY	1401	LysLysLysCysAspGluLeuAlaAlaAlaLeuArgGlyMetGlyLeuAsnSerValAla	1420	Db	5621	GTAGAGCAATTTCTGGGCCAACAACATGTGGAATCTTATCAGCGGCATTCATACCTCGCA	5680
Db	4541	AAGAAAGTGTGACGAGCTCGCGGGCGCCCTTCCGGGGTATGGGCTTGAACGAGTGGCA	4600	QY	1781	GlyLeuSerThrLeuProGlyAsnProAlaValAlaSerMetMetAlaPheSerAlaAla	1800
QY	1421	TyrTyrArgGlyLeuAspValSerValIleProThrGlnGlyAspValValValAla	1440	Db	5681	GGACTATCAACATGCGCAGGGAACCTCTGCTGTAGCTTCCATGATGGCATTCAGTGGCGCC	5740
Db	4601	TACTACAGAGGGCTGAGCGCTCTCGTAAATACCAACTCAGGAGACGTAGTGTGCTGCC	4660	QY	1801	LeuThrSerProLeuSerThrSerThrThrIleLeuLeuAsnIleLeuGlyGlyTyrLeu	1820
QY	1441	ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnVal	1460	Db	5741	CTCACCAGTCCGTTGTCAACTAGCACCTATCTTCTCAACATTTTGGGGGCTCGCTA	5800
Db	4661	ACCGAGCGCCCTCATGACGGGGTTTACTGGAGACTTTGACTCCGCTGATCGACTGCAACGTA	4720	QY	1821	AlaSerGlnIleAlaProProAlaGlyAlaThrGlyPheValValSerGlyLeuValGly	1840
QY	1461	AlaValThrGlnValAlaAspPheSerLeuAspProThrPheThrIleThrGlnIle	1480	Db	5801	GCATCCCAATTTGGCGCTCCCGGGGGCTACCGGCTTCTGTCTAGTGGCTTGGTGGG	5860
Db	4721	CGGGTCACTCAAGTGTGTAGACTTCACTTGGACCTTGGACCTTGGATCGACTGCAACGTA	4780	QY	1841	AlaAlaValGlySerIleGlyLeuGlyLysValLeuValAspIleLeuAlaGlyTyrGly	1860
QY	1481	ValProGlnAspAlaValSerArgSerGlnArgArgGlyThrGlyArgGlyArgLeu	1500	Db	5861	GCTGCCGTAGGCGAGCATAGGCTTGGGTAAGGTGCTGGTGGACATCTCTGGCAGGGTATGTT	5920
Db	4781	GTCCCTCAAGACGCTGTCTACGTAGCCAGCGCGGGCGCGCAGGCGGGGGAAGACTG	4840	QY	1861	AlaGlyIleSerGlyAlaLeuValAlaPheLysIleMetSerGlyGlyLysProSerMet	1880
QY	1501	GlyIleTyrArgTyrValSerThrGlyAlaSerGlyMetPheAspSerValVal	1520	Db	5921	GCGGGCATTTCCGGGGCTCTCTCGTCGATTTCAAGATCATCTGTGGCGAGAAGCCCTCCATG	5980
Db	4841	GGTATTATAGGTATGTTTCCATCTGTGTGAGCGAGCCTCAGGAATGTTTGACAGTGTAGT	4900				

QY	1881	GluAspValValAsnLeuLeuProGlyIleLeuSerProGlyAlaLeuValValGlyVal	1900	Db	7061	GTGACCCGATAGACTCTGAGTCCAAAGTGGTCTGAGACTCTCTCGACCAATGGTC	7120
Db	5981	GAGGATGTTGTCAACCTGCTGCTGGAATCTCTCTCCGGGTGCTGCTGGTGGAGTC	6040	QY	2261	GluGluAArgSerAspLeuGluProSerIleProSerGluTyrMetLeuProLysLysArg	2280
QY	1901	IleCysAlaAlaIleLeuArgArgHisValGlyProGlyGlyAlaValGlnTrpMet	1920	Db	7121	GAAGAAAGGAGCGACTTGGAGCTTCGATACCATCGGAATATATGCTCCCAAGAGAGA	7180
Db	6041	ATCTCGCGGCCATCTCTGGCCGACACAGTGGGACCGGGGAGGCGCTGTCCAATGGATG	6100	QY	2281	PheProAlaLeuProAlaTrpAlaArgProAspTyrAsnProProLeuValGluSer	2300
QY	1921	AsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaProThrHisTyrValThr	1940	Db	7181	TTCCCAACAGCCCTTACCGGCTTGGGACCGGCTGATTACAAACCCACCGCTTGTGGATCG	7240
Db	6101	AATAGGCTCATTCCTCTTGTCTCAGAGGAACACAGTGGCCGCCCACTACGTGACG	6160	QY	2301	TrpLysArgProAspTyrGlnProAlaThrValAlaGlyCysAlaLeuProProArg	2320
QY	1941	GluSerAspAlaSerGlnArgValThrGlnLeuLeuGlySerLeuThrIleThrSerLeu	1960	Db	7241	TGAGAGAGGCCAGATTACCAACCGGCACTGTGGGGGTGGCTCTCCCGCCCTAAG	7300
Db	6161	GAGTCGATCGTCGACGCGTGACCCCACTACTTGGCTCCCTTACCAATACCAACGCTG	6220	QY	2321	LysThrProThrProProArgArgArgArgThrValGlyLeuSerGluAspSerIle	2340
QY	1961	LeuArgArgLeuHisAsnTrpIleThrGluAspCysProIleProCysGlyGlySerTrp	1980	Db	7301	AAAAACCCGACGCTCCCAAGGAGACCGCGACAGTGGGTCTGAGTGAGAGCTCCATA	7360
Db	6221	CTCAGGAGACTCCACCACTGATTAAGAGCTGGCCCATCCATCGAGCGCTCGTGG	6280	QY	2341	GlyAspAlaLeuGlnLeuAlaIleLysSerPheGlyGlnProProProSerGlyAsp	2360
QY	1981	LeuArgAspValTrpAspTrpValCysThrIleLeuThrAspPheLysAsnTrpLeuThr	2000	Db	7361	GCAGATGCCCTAACAGCTGGCCATCAAGTCTTTGGCCAGCCCCCCCCCAAGCGCGAT	7420
Db	6281	CTCGCGATGTGGGATGGTGTGACCATCTAACAGACTTTAAAAAATGGCTGACC	6340	QY	2361	SerGlyLeuSerThrGlyAlaGlyAlaAlaAspSerGlySerGlnThrProProAspGlu	2380
QY	2001	SerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTyrLys	2020	Db	7421	TCAGGCTTTTCCACGGGGCGGACGACGCGATTCGGCGAGTCGGACGCCCCCGATGAG	7480
Db	6341	TCCAAAATGTTCCCAAGATGCTGCTGCTCCCTTATCTTGTCAAAAGGGGTACAAG	6400	QY	2381	LeuAlaLeuSerGluThrGlySerIleSerSerMetProProLeuGluGlyLeuLeuGly	2400
QY	2021	GlyValTrpAlaGlyThrGlyIleMetThrArgCysProCysGlyAlaAsnIleSer	2040	Db	7481	TTGGCCCTTTCGAGACAGTTCATCTCTCCATGCCCCCTCTCGAGGGGGAGCTGGA	7540
Db	6401	GGCGTGTGGCTGGCACTGATCATGACACACAGCTGCTCTTGGCGGCCAATATCTCT	6460	QY	2401	AspProAspLeuGluProGlnValGluProGlnProProGlnGlyGlyValAla	2420
QY	2041	GlyAsnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTrp	2060	Db	7541	GATCCAGACTTGAGGCTGAGCAGGTAGAGCTTCAACCTCCCCCGAGGGGGTGGTA	7600
Db	6461	GGCAATGTCGCTGGCTCATGAGAAATACGGGGCCCCAAAACCTGCATGAATATCTGG	6520	QY	2421	AlaProGlySerAspSerGlySerTrpSerThrCysSerGluLysAspSerValVal	2440
QY	2061	GlnGlyThrPheProIleAsnCysTyrThrGluGlyGlnCysValProLysProAlaPro	2080	Db	7601	ACCCCGGCTCAGGCTCGGGCTTGGTCTTACTTGTCTCCGAGGAGGACGACTCGTCTG	7660
Db	6521	CAGGGAGACCTTCCCATCAATGTTACAGGAGGGCGGTGCGTGGCGAAACCCGACCA	6580	QY	2441	CysCysSerMetSerTyrSerTrpThrGlyAlaLeuIleThrProCysSerProGluGlu	2460
QY	2081	AsnPheLysValAlaIleTrpArgValAlaAlaSerGluTyrAlaGluValThrGlnHis	2100	Db	7661	TGCTGTCTCATGTGCTACTCTCGACCCGGGGCTCTAATAACTCTTGTAGCCCCGAGAG	7720
Db	6581	AACTTTAAGATCGCCATCTGAGGGTGGCGGCTCAGAGTACGCGGAGGTGACGACAC	6640	QY	2461	GluLysLeuProIleAsnProLeuSerAsnSerLeuLeuArgTyrHisAsnLysValTyr	2480
QY	2101	GlySerTyrHisTyrIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeu	2120	Db	7721	GAAGAATGGCAATTTGGCCCTTGGCAACTCCCTGTTGGGATATCACCAAGGTGTAC	7780
Db	6641	GGGTATACCATACATACAGGACTTACCACTGATTAAGTTCCTTGGCAACTA	6700	QY	2481	CysThrThrThrLysSerAlaSerLeuArgAlaLysLysValThrPheAspArgMetGln	2500
QY	2121	ProSerProGluPhePheSerTrpValAspGlyValGlnIleHisArgPheAlaProThr	2140	Db	7781	TGTACCAATCAAGAGCGCTCATTAAGGGCTAAAGAGTAACCTTTGTAGATGATGCAA	7840
Db	6701	CCTTCTCCAGATTCCTTTCTGGGTGGAGCGAGTGCAGATTCATAGGTTTGGCCCCATA	6760	QY	2501	ValLeuAspSerTyrTyrAspSerValLeuLysAspIleLysLeuAlaSerLysVal	2520
QY	2141	ProLysProPhePheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal	2160	Db	7841	GCCTTCGAGCGCTCATTTAGCTCAGTCTTGAAGGACATTAAGCTAGCGGCTCCAAAGTC	7900
Db	6761	CCGAGCGGTTTTTCGGGATAGGTCTCGTCTGCGTGGGCTTAATTCATTTGCTGTC	6820	QY	2521	ThrAlaArgLeuThrMetGluGluAlaCysGlnLeuThrProProHisSerAlaArg	2540
QY	2161	GlySerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThr	2180	Db	7901	ACCGAAGGCTTCTCACTTTAGAGAGGCTCGCAGTTAACTCCACCCCACTCTGCNAGA	7960
Db	6821	GGGTCTCAGCTCCCTTGCATCTCTGAACCTGACACAGACGATTAATGACGTCATGCTAACA	6880	QY	2541	SerLysTyrGlyPheGlyAlaLysGluValArgSerLeuSerGlyArgAlaValAsnHis	2560
QY	2181	AspProSerHisIleThrAlaGluThrAlaAlaArgArgLeuAlaArgGlySerProPro	2200	Db	7961	TCCAAGTATGGTTTGGGGCTAAGGAGTCCGAGCTTGTCCGGGAGAGCGGTTAACCCAC	8020
Db	6881	GACCATCCCATATACGCGGAGACTGACGCGGGGCTTGGACCGGGGGTCAACCCCG	6940	QY	2561	IleLysSerValTrpLysAspLeuLeuAspSerGluThrProIleProThrIle	2580
QY	2201	SerGluAlaSerSerAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThr	2220	Db	8021	ATCAAGTCCGTGTGGAGGACCTCTCTGGAAGACACACAAACACCAATTCCTACCAACATC	8080
Db	6941	TCCAGGACAGCTCTCTCAGCGAGCGAGTATCGCACCATCGCTGCGAGCCACCTGCACC	7000	QY	2581	MetAlaLysAsnGluValPheCysValAspProThrLysGlyGlyLysLysAlaAlaArg	2600
QY	2221	ThrHisGlyLysAlaTyrAspValAspMetValAspAlaAsnLeuPheMetGlyGlyAsp	2240	Db	8081	ATGCCCAAAATAGGTGTTCTCGGTGGACCCCAACAGGGGGGTAAAGAGGAGCTCGC	8140
Db	7001	ACCCACGCAAGGCTATGATGTGGACATGGTGGATGCCAACCTGTTTCATGGGGGGCGAT	7060	QY	2601	LeuIleValTyrProAspLeuGlyValArgValCysGluLysMetAlaLeuTyrAspIle	2620
QY	2241	ValThrArgIleGluSerGlySerLysValValValLeuAspSerLeuAspProMetVal	2260				

Db 8141 CTTATCGTTTACCTGACCTCGCGCTCAGGGTCTCGAGAGAAATGGCCCTTTATGATATC 8200
 Qy 2621 ThrGlnLysLeuProGlnAlaValMetGlyAlaSerTyrGlyPheGlnTyrSerProAla 2640
 Db 8201 ACACAAAGCTTCTCAGGGCGGTGATGGGGCTTCTTATGATTCAGTACTCCCGGCT 8260
 Qy 2641 GlnArgValGluPheLeuLeuLysAlaTrrPAlaGluLysLysPheProMetGlyPheSer 2660
 Db 8261 CAGCGGTGGAGTTCTCTTGAAGGATGCGCGGAAAGAACACCTATGGGTTTTCG 8320
 Qy 2661 TyrAspThrArgCysPheAspSerThrValThrGluArgAspLeuThrGluSer 2680
 Db 8321 TATGATACCGATGCTTTGACTCAACCGTCACTGAGAGACATCAGGATGAGGATCC 8380
 Qy 2681 IleTyrArgAlaCysSerLeuProGluGluAlaHisThrAlaIleHisSerLeuThrGlu 2700
 Db 8381 ATATATCGGCTTGTCTTCCCGAGGAGGCCACACATGCCATACACTCACTGACTGAG 8440
 Qy 2701 ArgLeuTyrValGlyGlyProMetPheAsnSerLysGlyGlnThrCysGlyTyrArgArg 2720
 Db 8441 AGACTTTACGTGGAGGGCCCATGTTCAACAGCAAGGGCCAGACCTCGGGGTACAGCGT 8500
 Qy 2721 CysArgAlaSerGlyValLeuThrThrSerMetGlyAsnThrIleThrCysTyrValLys 2740
 Db 8501 TGGCGGCCAGCGGGGTGCTTACACTAGCATGGGGAACACCATCATGCTATGTGAA 8560
 Qy 2741 AlaLeuAlaAlaCysLysAlaAlaGlyIleAlaProThrMetLeuValCysGlyAsp 2760
 Db 8561 GCCTTAGCGGCTGTAAAGCTGAGGATTAATTCGCCCCACATGCTGTATGGGCGAT 8620
 Qy 2761 AspLeuValValLysSerGluSerGlnGlyThrGluGluAspGluArgAsnLeuArgAla 2780
 Db 8621 GACTTGTGTCATCTCAGAGAGCCAGGGGACCGAGGAGGAGCGGAACCTGAGAGCC 8680
 Qy 2781 PheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProArgProGluTyr 2800
 Db 8681 TTCAGGAGGCTATGACAGGTATTCGCCCCCTCTGTGTGACCCCGCCAGCCGGAATAT 8740
 Qy 2801 AspLeuGluLeuIleThrSerCysSerAsnValSerValAlaLeuGlyProGlnGly 2820
 Db 8741 GACTGGAGCTGATNAACCTTCTCTCAATGTCTGTGTGGCGTGGGCCCCACAGGC 8800
 Qy 2821 ArgArgArgTyrTyrLeuThrArgAspProThrThrProIleAlaArgAlaAlaTrrPAla 2840
 Db 8801 CGCGCGAGATACCTGACAGAGACCTTACCACCTCCATCGCCCGGGCTGCTGGAA 8860
 Qy 2841 ThrValArgHisSerProValAsnSerTrrPLeuGlyAsnIleIleGlnTyrAlaProThr 2860
 Db 8861 ACAGTTAGACACTCCCGCTGTCAATTCATGCTTAGGAACATCATCCAGTACGCCCAACC 8920
 Qy 2861 IleTrrPAlaArgMetValLeuMetThrHisPhePheSerIleLeuMetAlaGlnAspThr 2880
 Db 8921 ATATGGCTCGCATGTCTGTGATGACACATTTCTTCCATTTCTATGGCCCAAGATACT 8980
 Qy 2881 LeuAspGlnAsnLeuAsnPheGluMetTyrGlyAlaValTyrSerValSerProLeuAsp 2900
 Db 8981 CTGGCCAGAACCTCAACTTTGAGATGATCGAGCGGTGTACTCCGTGAGTCCCTTGGAC 9040
 Qy 2901 LeuProAlaIleIleGluArgLeuHisGlyLeuAspAlaPheSerLeuHisThrTyrThr 2920
 Db 9041 CTCCAGCCATAATGAAGGTTTACAGGGCTTTCAGCGCTTTCCTCTCTGACACATACACT 9100
 Qy 2921 ProHisGluLeuThrArgValAlaSerAlaLeuArgLysLeuGlyAlaProProLeuArg 2940
 Db 9101 CCCACGAACTGACAGGGTGGCTTCAGCCCTCAGAAACTTGGGGCGCCACCCCTCAGA 9160
 Qy 2941 AlaTrrPysSerArgAlaArgAlaValArgAlaSerLeuIleSerArgGlyGlyArgAla 2960
 Db 9161 GCGTGAAGAGCGGGACCTGTCAGTCAGTCCAGGCGTCCCTCATCTCCCGTGGGGGAGCG 9220
 Qy 2961 AlaValCysGlyArgTyrLeuPheAsnTrrPAlaValLysThrLysLeuThrPro 2980
 Db 9221 GCCGTTTGGCGCGATATCTTCTTCACTGGGCGGTGAAGACCAAGCTCAAACTCACTCCA 9280

Qy 2981 LeuProGluAlaArgLeuLeuAspLeuSerSerTrrPheThrValGlyAlaGlyGlyGly 3000
 Db 9281 TTGCGGGAAGCGGCTCTCTGGATTTATCAGCTGTTTCACTGTCGCGCGCGGGGC 9340
 Qy 3001 AspIleTyrHisSerValSerArgAlaArgProArgLeuLeuPheGlyLeuLeu 3020
 Db 9341 GACATTTATCACAGCGTGTGCGTGCAGCCCGCTTATTACTCTTGGCTACTCCTA 9400
 Qy 3021 LeuPheValGlyValGlyLeuPheLeuLeuProAlaArg 3033
 Db 9401 CTTTTGTAGGGGTAGCCCTTTTCTACTCCCGCTCGG 9439
 RESULT 7
 E07362
 LOCUS CDNA of Hepatitis non-A non-B virus. linear PAT 29-SEP-1997
 DEFINITION E07362
 ACCESSION E07362.1 GI:2175501
 VERSION JP 1994121689-A/2.
 KEYWORDS
 SOURCE unidentified
 ORGANISM
 unclassified.
 REFERENCE 1 (bases 1 to 9589)
 AUTHORS Okamoto, H. and Nakamura, T.
 TITLE NON-A NON-B HEPATITIC VIRUS GENE, POLYNUCLEOTIDE, POLYPEPTIDE,
 ANTIGEN AND ANTIBODY DETECTION SYSTEM
 JOURNAL Patent: JP 1994121689-A 2 06-MAY-1994;
 NAKAMURA TETSUO
 COMMENT OS Hepatitis non-A non-B virus
 PN JP 1994121689-A/2
 PD 06-MAY-1994
 PF 09-AUG-1991 JP 1991287402
 PI OKAMOTO HIROAKI, NAKAMURA TETSUO
 PC C12N15/51, A61K39/29, A61K39/395, A61K39/395, C07K13/00, C12P21/02,
 C12P21/08,
 PC G01N33/53, G01N33/576, G01N33/577//A61B10/00;
 CC strandedness: Single;
 CC topology: Linear;
 CC hypothetical: No;
 CC anti-sense: No; Location/Qualifiers
 FH Key
 FH
 FT source 1..9589
 FT /organism='Hepatitis non-A non-B virus' FT
 FT /strain='HCV-J6'.
 FEATURES
 source Location/Qualifiers
 1..9589
 /organism='unidentified'
 /mol_type='genomic RNA'
 /db_xref='taxon:3264'
 BASE COUNT 1968 a 2820 c 2635 g 2166 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 9589
 Score: 15824.00 Matches: 2966
 Percent Similarity: 98.78% Conservative: 30
 Best Local Similarity: 97.79% Mismatches: 37
 Query Match: 98.16% Indels: 0
 DB: 6 Gaps: 0
 US-09-980-559-2 (1-3033) x E07362 (1-9589)
 Qy 1 MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGln 20
 Db 341 ATGAGCACAAATCTTAAACCTCAAGAGAAACCAAGAAACCAACCAACCGTCCGCCAAC 400
 Qy 21 AspValLysPheProGlyGlyGlyGlnIleValGlyValTrrLeuLeuProArgArg 40
 Db 401 GACGTTAAGTTTCCGGCGCGCGCCAGATCGTTGGCGGAGTATACTTGTTCGCCCGCAGG 460
 Qy 41 GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly 60

1541	GGCATGTTCTCCCTTGGTGGCCAGGAGNAANAATCCAGCTCATCAACACCAATGGCGAGTTGG	1600
421	HisIleAenAArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer	440
1601	CACATCAACCGCACCGCCCTGAATGCAATGACATCTTTTGCAACACCGGCTTCTCGCGTCA	1660
441	LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluuAArgMetSerAlaCysArg	460
1661	CTGTTTACACCCACAGCTTCAACTCGTFCAGAGTGTCCCGAACGATGTCCGCTCGCGC	1720
461	SerIleGluAlaPheArgValGlyTyrGlyAlaLeuGlnTyrGluAspAsnValThrAsn	480
1721	AGTATCGAGGCTTTTCGGGTGGAGTGGGGCGCTTACAAATAGAGACAATGTACCACAT	1780
481	ProGluAspMetArgProTyrCysTrpHisTyrProProArgGlnCysGlyValValSer	500
1781	CCAGAGGATATGAGACCGTATTGTGGCACTACCCACCAAGACAGTGTGGTAGTCTCC	1840
501	AlaIysThrValCysGlyProValTyrCysPheThrProSerProValValValGlyThr	520
1841	CGGAGCTCTGTGTGGCCAGTACTGTGTTTACCCCGACGCCCATAGTAGTGGGTAGC	1900
521	ThrAspArgLeuGlyAlaProThrTyrThrTyrGlyGluAsnGluThrAspValPheLeu	540
1901	ACCGATAGCTTGGAGCGCCCACTTACACGTGGGGGGAGAAATGACACAGATGTCTTCTTA	1960
541	LeuAsnSerThrArgProProLeuGlySerTrpPheGlyCysThrTrpMetAsnSerSer	560
1961	TTGAACAGCACTCGNACCACCGCAGGGGTTCATGGTTTCGGCTGCACGTGGATGAATCCACT	2020
561	GlyTyrThrIysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer	580
2021	GGCTACCAAGACTTGGCGGGCCACACCCCTGCCCATTTAGAGCTGTACTTCAATGCCAGC	2080
581	ThrAspLeuLeuCysProThrAspCysPheArgIysHisProAspThrThrTyrLeuIys	600
2081	ATGGACTTGTGTGCCCCCAGCGACTGTTTTAGGAAGCATCTTGATACCACTTACATCAAA	2140
601	CysGlySerGlyProTyrLeuThrProArgCysLeuIleAspTyrProTyrArgLeuTyr	620
2141	TGTGGCTTGGGCCCTGGCTCAGCCAGGTGCCTGATCGACTACCCCTACAGGCTCTGG	2200
621	HisTyrProCysThrValAsnTyrThrIlePheIysIleArgMetTyrValGlyVal	640
2201	CATTATCCCTGCACAGTTAACTATACCATCTTCAAAATAAGGATGTATGTGGGGGGTCTC	2260
641	GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp	660
2261	GAGCACAGCTCACCGCTCGGTGCAATTTCACTCGTGGGGATCGTGTCAACTTTGGAGGAC	2320
661	ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGluTyrAlaIleLeuPro	680
2321	AGAGACAGAGTCAACTGTCTCTTGTGTCACCTCCACACGAGGTGGGCCATTTTACCT	2380
681	CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisGlnAsnIle	700
2381	TGCACCTTACTTGGACCTGCCCGCTTGTGCAGCTGGTCTTCTCCACCTCCACCAAAACATC	2440
701	ValAspValGlnPheMetTyrGlyIleuSerProAlaLeuThrIysTyrIleValArgTyr	720
2441	GTGGAGCGTCAATTCATGTATGGCCCTATCACCTGTCTCACAAAATACATCGTCCGATGG	2500
721	GluTyrValIleLeuLeuPheLeuLeuAlaAspAlaArgValCysAlaCysLeuTyr	740
2501	GAGTGGGTAGTACTCTTATTCCTGTCTTATAGCGGACCGCAGGGTTTGGCTTGTATGG	2560
741	MetLeuIleLeuLeuGlyGlnAlaGluAlaLeuGluIysLeuValIleLeuHisAla	760
2561	ATGCTCATCTTGTGGCCAGGCCGGAAGCAGCACTAGAGAAATGTGTGTCTTGTTCACGCT	2620
761	AlaSerAlaAlaSerCysAsnGlyPheLeuTyrPheValIlePhePheValAlaAlaTyr	780
2621	GGAGCGCAGCTAGCTGCAATGGCTTCTTATACTTTGTCTCATCTTTTTCGTGGCTGTCTGG	2680

QY	781	TyrIleLysGlyArgValProLeuAlaThrTyrSerLeuThrGlyLeuTrpSerPhe	800
DB	2681	TACATCAAGGGTCGGGTAGTCCCTTGGCTACTATTTCCTCTCACTGGCTTATGGTCTCTTT	2740
QY	801	SerLeuLeuLeuAlaLeuProGlnAlaTyrAlaTyrAspAlaSerValHisGly	820
DB	2741	GGCTTACTGCTCTCTAGCATTTGCCCAACAGGCTTATGCTTATACGCATCTGTACATGGT	2800
QY	821	GlnIleGlyAlaAlaLeuLeuValMetIleThrLeuPheThrLeuThrProGlyTyrLys	840
DB	2801	CAGATAGGACAGCTCTGTTGGTACTGATCACTCTCTTTACATCTACCCCGGGTATAG	2860
QY	841	ThrLeuLeuSerArgPheLeuTrpTrpLeuCysTyrLeuLeuThrLeuGlyGluAlaMet	860
DB	2861	ACCCTTCTCAGCCGGTTCTGTGGTGTGTCTATCTTCTGACCTCGCGGAAGCTATG	2920
QY	861	ValGlnGluTrpAlaProMetGlnValArgGlyGlyArgAspGlyIleIleTrpAla	880
DB	2921	GTCAGGAGTGGGCACCACTATGCAGGTGCGGGTGGCGGTGATGGATCATATGGGCC	2980
QY	881	ValAlaIlePheTyrProGlyValValPheAspIleThrIleTrpLeuLeuAlaValLeu	900
DB	2981	GTCCCATATCTTCCCGGGTGGTGTGGTGGATTAACCAAGTGGCTCTTGGCGGTGCTT	3040
QY	901	GlyProAlaTyrLeuLeuLysGlyAlaLeuThrArgValProTyrPheValArgAlaHis	920
DB	3041	GGGCTGCTTATCTCTCTAAAGGTGCTTGCAGCGGTGGCGTACTTCTGCAGGGCTCAC	3100
QY	921	AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyArgTyrValGlnMet	940
DB	3101	GCTCTACTAAGGATGTGCACCATGTAAGCATCTCGCGGGGTAGGTACGTCCAGATG	3160
QY	941	AlaLeuLeuAlaLeuGlyArgTyrThrGlyThrTyrIleTyrAspHisLeuThrProMet	960
DB	3161	GTGCTACTAGCCCTTGGCAGGTGGACTGGGCATTTACATCTATGACCACCTTCACCCCTATG	3220
QY	961	SerAspTrpAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIleIlePhe	980
DB	3221	TCGGATTGGCTGCTAATGSCCTCGGGACTTGGCGGTGGCGGTGAGCTTATCATCTTC	3280
QY	981	SerProMetGlnLysLysValIleValTrpGlyAlaGluThrAlaAlaCysGlyAspIle	1000
DB	3281	AGTCCGATGGAGAAAAAGTCATCGTCTGGGAGCGGAGACAGCTGCTTGGCGGGATATC	3340
QY	1001	LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp	1020
DB	3341	TTACACGGACTTCCCGGTCTCCGCCGACTTGGCGGGAGGTCTCTTGGCCCGAGCTGAT	3400
QY	1021	GlyTyrThrSerLysGlyTrpSerLeuLeuAlaProIleThrAlaTyrAlaGlnGlnThr	1040
DB	3401	GGCTATACCTCCAAAGGGGTGAGTCTTCTGCCGCCCATCACTGCTTATGCCACGACACA	3460
QY	1041	ArgGlyLeuLeuGlyThrIleValSerMetThrGlyArgAspLysThrGluGlnAla	1060
DB	3461	CGCGGGCTTTGGGCACCATAGTGGTGAGCATGACGGGGCGGACAGAAGCAGGCC	3520
QY	1061	GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly	1080
DB	3521	GGGGAGATTTCAGTCTGCTCCACGGTCACTCACTGCTCTCTCGGAAACAACCATCTCGGG	3580
QY	1081	ValLeuTrpThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro	1100
DB	3581	GTCTTATGAGTGTCTACCATGAGCTGGCAACAGACTCTACCGGGTGGCCAGCCCGGGTCCG	3640
QY	1101	ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTrpProSerProGly	1120
DB	3641	GTCAACAGATGTACTCCAGTGTGAGGGGACTTATGGGGTGGCGGCGCCAGCCCGGG	3700
QY	1121	ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrIleValThrArgAsn	1140
DB	3701	ACCAATATCTTTGGAGCGGTGCAGTGTGGAGCGGTGCAGCTATACCTGCTGTCAGCGAAAC	3760
QY	1141	AlaAspValIleProAlaArgArgArgGlyAspLysArgGlyAlaLeuLeuSerProArg	1160
DB	3761	GCTGATGTCTCCCGCTCGAAGACGCGGGGACAAGCGAGGAGCGCTACTCTCCCCGAGA	3820
QY	1161	ProLeuSerThrLeuLysGlySerSerGlyGlyProValLeuLeuCysProArgGlyHisAla	1180
DB	3821	CCTCTTTCCACCTTGAAGGGGTCTCTGGGGGGCCCGGTCTCTGCCCCAGAGCCACGCT	3880
QY	1181	ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle	1200
DB	3881	GTCCGGGTCTTCCGGGCACCGGTGTCTCCGGGGGTGGCCAGTCCATAGATTTATC	3940
QY	1201	ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro	1220
DB	3941	CCCGTTGAGACACTTGACATCGTCACTCGGTCCCCCACCCTTTAGTGACAACAGCACCA	4000
QY	1221	ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLys	1240
DB	4001	CCTGCTGTGCCCAACTTATCAGGTCTGGGTACTTACATGCCCGGCTGGTAGTGGAAAG	4060
QY	1241	SerThrLysValProValAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnPro	1260
DB	4061	AGCACCAAGTCCCTGTCTCGGTATGCCGTCAAGGGGTACAAGGTGTAGTGTATATCCC	4120
QY	1261	SerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerLysAlaHisGlyIleAsnPro	1280
DB	4121	TCGGTGGCTGCCACCTCGGGTTTGGGGGTACTTGTCCCAAGGCACATGGCATCAATCCC	4180
QY	1281	AsnIleArgThrGlyValArgThrValThrGlyAlaProIleThrTyrSerThrTyr	1300
DB	4181	AACATTTAGACTTGGGTCAAGGACTGTGACACCGGGGGCGCCCATCATCGTACTCCATAT	4240
QY	1301	GlyLysPheLeuAlaAspGlyGlyCysAlaGlyGlyValaTyrAspIleIleLysCysAsp	1320
DB	4241	GGCAAAATTCCTGCCGATGGGGGTGCGCAGCGGGCGGCTATGACATCATATTCGGAT	4300
QY	1321	GluCysHisAlaValAspSerThrIleLeuGlyIleGlyThrValLeuAspGlnAla	1340
DB	4301	GAATGCCATGCGGTGGACTCTACCACTTCTCGGCATCGGAACAGTCTCGATCAAGCA	4360
QY	1341	GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProProGlySerValThr	1360
DB	4361	GAGACAGCGGGGTCAAGGCTAACTGTACTGGGTACGGCTACGCCCGCCCGGGTCAGTGACA	4420
QY	1361	ThrProHisProAsnIleGluGluValAlaLeuGlyGlnGlyGluIleProPheTyr	1380
DB	4421	ACCCGCCACCCCAACATAGAGAGGTGGCCCTCGGCAGAGGTGAGATCCCTCTCTAT	4480
QY	1381	GlyArgAlaIleProLeuSerTyrIleLysGlyArgHisLeuIlePheCysHisSer	1400
DB	4481	GGGAGGGCGATTCCCTGTCTACATCAAGGAGGGAAGACACTTGATCTTCTGCCACTCA	4540
QY	1401	LysLysLysCysAspGluLeuAlaAlaLeuArgGlyMetGlyLeuAsnSerValAla	1420
DB	4541	AAGAAAAAGTGTACAGACTCGCGCGGCCCTTCGGGGTATGGGCTTGAACGAGTGCGCA	4600
QY	1421	TyrTyrArgGlyLeuAspValSerValIleProThrGlnGlyAspValValValAla	1440
DB	4601	TACTACAGAGGGTGGACGCTCCCGTAATACCACTCAGGAGAGCGTAGTGGTCTGGCC	4660
QY	1441	ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnVal	1460
DB	4661	ACCGAGCCCTCATGCGGGGTTTACTGGAGACTTTTGACTCCGTGATCCGACTGCAAGTA	4720
QY	1461	AlaValThrGlnValValAspPheSerLeuAspProThrPheThrIleThrThrGlnIle	1480
DB	4721	GGGGTCACTCAAGTGTAGACTTCAGCTTGGAGCCCACTTCCACATAACACACAGACT	4780
QY	1481	ValProGlnAspAlaValSerArgSerGlnArgGlyValArgGlyValArgGlyVal	1500
DB	4781	GTCCCTCAAGACGCTGTCTCAGTAGCCGCGGGCGGCGGAGGAGGAGACTG	4840
QY	1501	GlyIleTyrArgTyrValSerThrGlyGluArgAlaSerGlyMetPheAspSerValVal	1520

QY	2241	ValThrArgIleGluSerGlySerLysValValValLeuAspSerLeuAspProMetVal	2260	QY	2601	LeuIleValTyrProAspLeuGlyValArgValCysGluLysMetAlaLeuTyrAspIle	2620
DB	7061	GTGACCCGGATAGAGTCTGAGTCCAAAGTGGTCTCTGGACTCTCTCGACCCCAATGGTTC	7120	DB	8141	CTTATCGTTTACCTTGACCTCGCGCTCGAGGCTCGAGAAATGSCCTTTATGATATC	8200
QY	2261	GluGluArgSerAspLeuGluProSerIleProSerGluTyrMetLeuProLysLysArg	2280	QY	2621	ThrGlnLysLeuProGlnAlaValMetGlyAlaSerTyrGlyPheGlnTyrSerProAla	2640
DB	7121	GAAGAAAGGAGGACCTTGAGCTTCGATACCATCGGAATATATGCTCCCAAGAAGAGA	7180	DB	8201	ACACAAAGCTTCTCTCAGCGGTGATGGGGCTTCTTATGGATTCCAGTACTCCCGCT	8260
QY	2281	PheProAlaLeuProAlaTyrAlaAatgProAspTyrAsnProProLeuValGluSer	2300	QY	2641	GlnArgValGluPheLeuLeuLysAlaTyrAlaGluLysLysAspProMetGlyPheSer	2660
DB	7181	TTCCACACCGCTTACCGCTTGGGACCGCCCTGATTTACAACCCACCGCTTGGGAATCG	7240	DB	8261	CAGCGGTGGAGTTCTCTTGAGGCTGGGGGAGGAGAAAGAACCCCTATGGGTTTTTCG	8320
QY	2301	TyrLysArgProAspTyrGlnProAlaThrValAlaGlyCysAlaLeuProProArg	2320	QY	2661	TyrAspThrArgCysPheAspSerThrValThrGluArgAspIleArgThrGluGluSer	2680
DB	7241	TGGAAGAGGCCAGATTACCAACCGCCACTGTTGGCGGCTGGCTCTCCCGCCCTTAAG	7300	DB	8321	TATGATACCCGATGCTTTGACTCAACCGTCACTGAGAGAGACATCAGGACTGAGGAGTCC	8380
QY	2321	LysThrProThrProProProArgArgArgThrValGlyLeuSerGluAspSerIle	2340	QY	2681	IleTyrArgAlaCysSerLeuProGluGluAlaHisThrAlaIleHisSerLeuThrGlu	2700
DB	7301	AAAACCCGACGCTCCCGACGAGACGCGGACAGTGGGTCTGAGTGAGAGCTCCATA	7360	DB	8381	ATATATCGGGCTTGTTCCTTGGCCGAGAGGCCACACTGCCATACACTCACTGACTGAG	8440
QY	2341	GlyAspAlaLeuGlnGlnLeuAlaIleLysSerPheGlyGlnProProProSerGlyAsp	2360	QY	2701	ArgLeuTyrValGlyGlyProMetPheAsnSerLysGlyGlnThrCysGlyTyrArgArg	2720
DB	7361	GCAGATGCCCTCAACAGCTGGCCATCAAGTCTTTGGCCAGCGCCCGCCCAAGCGCAT	7420	DB	8441	AGACTTTACGTGGAGGGCCCATGTTCAACAGCAAGGGCCAGACTGGCGGTACAGCGT	8500
QY	2361	SerGlyLeuSerThrGlyAlaGlyAlaAlaAspSerGlySerGlnThrProProAspGlu	2380	QY	2721	CysArgAlaSerGlyValLeuThrThrSerMetGlyAsnThrIleThrCysTyrValLys	2740
DB	7421	TCAGGCTTTCCACGGGGCGGACGCGCCGATTCGGCAGTCGGACGCCCGCCCGATGAG	7480	DB	8501	TGCGCGCCAGCGGGGTGCTTACCACTAGCATGGGGAACACCATCATGCTATGTGAAA	8560
QY	2381	LeuAlaLeuSerGluThrGlySerIleSerSerMetProProLeuGluGlyLeuGly	2400	QY	2741	AlaLeuAlaAlaCysLysAlaAlaGlyIleIleAlaProThrMetLeuValCysGlyAsp	2760
DB	7481	TTGGCCTTTCCGAGACAGCTTCCATCTCTCCATGCCCCCTCTCGAGGGGGAGCCTGGA	7540	DB	8561	GCCTTAGCGGCTGTAAAGCTCGAGGATAATTTGGCCGCCACAATGCTGCTATGCGCGAT	8620
QY	2401	AspProAspLeuProGluGlnValGluProGlnProProProGlnGlyValAla	2420	QY	2761	AspLeuValValIleSerGluSerGlnGlyThrGluGluAspGluArgAsnLeuArgAla	2780
DB	7541	GATCCAGACTGAGAGCTTGAGCAGGTAGAGCTTCAACCTCCCGCCCGAGGGGGGTGGA	7600	DB	8621	GACTTGGTGTCTATCTCAGAGGCCAGGGGACCGAGAGGACGAGGGAACTTGAGAGCC	8680
QY	2421	AlaProGlySerAspSerGlySerTyrProSerThrCysSerGluGluAspSerValVal	2440	QY	2781	PheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProArgProGluTyr	2800
DB	7601	ACCCCGGCTCAGGCTCGGGCTCTTGGTCTACTTGTCCGAGGAGGACGACTCCGCTCGT	7660	DB	8681	TTCCAGGAGCTATGACCAAGTATCTGCCCCCTCTGCTGCTGACCCCCCAGACCGGAATAT	8740
QY	2441	CysCysSerMetSerTyrSerTyrThrGlyAlaLeuIleThrProCysSerProGluGlu	2460	QY	2801	AspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaLeuGlyProGlnGly	2820
DB	7661	TGCTGCTCCATGTATCATCTCTGGACCGGGCTCTAATACTCCTTGTAGCCCCGAGAG	7720	DB	8741	GACTTGGAGCTGATAACATCTTGTCTCTCAATATGCTGTGTGGCGTGGGCCCCACAGGC	8800
QY	2461	GluLysLeuProIleAsnProLeuSerAsnSerLeuLeuArgTyrHisAsnLysValTyr	2480	QY	2821	ArgArgArgTyrTyrLeuThrArgAspProThrThrProIleAlaArgAlaTyrGlu	2840
DB	7721	GAAGAAGTTGCCAATTGGCCCTTGAGCAACTCTCTGTCGATATCACAAAGGTGTAC	7780	DB	8801	CGCCGAGAGATCTTACCTGACAGAGACCTTACCACTCCCAATCGCCGGGCTGCTGGAA	8860
QY	2481	CysThrThrThrLysSerAlaSerLeuAlaLysLysValThrPheAspArgMetGln	2500	QY	2841	ThrValArgHisSerProValAsnSerTrpLeuGlyAsnIleIleGlnTyrAlaProThr	2860
DB	7781	TGTACCAATCAAGAGCGCTCATTAAGGGCTTAAAGGTAACTTTGTATGATGTCAA	7840	DB	8861	ACAGTTAGACACTCCCTCTCAATTCATGCTAGGAAACATCATCCAGTACGCCCAACC	8920
QY	2501	ValLeuAspSerTyrTyrAspSerValLeuLysAspIleLysLeuAlaAlaSerLysVal	2520	QY	2861	IleTyrAlaArgMetValLeuMetThrHisPheSerIleLeuMetAlaGlnAspThr	2880
DB	7841	CGGCTCGAGCTCATTTATGACTCAGTCTTGAAGACATTAAGCTAGCGGCTCCAAGTTC	7900	DB	8921	ATATGGGCTCGATGCTGCTGATGACACACTTCTTCTCCATTTCTCAATGCCCCAAGTACT	8980
QY	2521	ThrAlaArgLeuThrMetGluAlaCysGlnLeuThrProProHisSerAlaArg	2540	QY	2881	LeuAspGlnAsnLeuAsnPheGluMetTyrGlyAlaValTyrSerValSerProLeuAsp	2900
DB	7901	ACCGCAAGGCTTCTCACTTAGAGGAGGCTCGCAGTTTAACTCCACCCCACTCTGAAGA	7960	DB	8981	CTGGACCAAGACTCAACTTTGAGATGACGAGCGGTGACTCCGCTGAGTCTCTGGAC	9040
QY	2541	SerLysTyrGlyPheGlyAlaLysGluValArgSerLeuSerGlyArgAlaValAsnHis	2560	QY	2901	LeuProAlaIleIleGluArgLeuHisGlyLeuAspAlaPheSerLeuHisThrTyrThr	2920
DB	7961	TCCAGTATGGGTTTGGGGCTTAAGAGGTCCGAGCTTGTCCGGGAGAGCCGTTAACCC	8020	DB	9041	CTCCAGCCCATTAATGAAGGTTACACGGGCTTGAGGCTTTCTCTCTGCACACATACACT	9100
QY	2561	IleLysSerValTyrLysAspLeuLeuGluAspSerGluThrProIleProThrIle	2580	QY	2921	ProHisGluLeuThrArgValAlaSerAlaLeuArgLysLeuGlyAlaProProLeuArg	2940
DB	8021	ATCAAGTCCGTGTGAAGGACCTCTTGAAGACACACAAACCAATTCCTACAAACATC	8080	DB	9101	CCCCACCAACTGACACGGGTGCTTCAAGCCCTCAGAAACTTGGGGCGCCACCCCTCAGA	9160
QY	2581	MetAlaLysAsnGluValPheCysValAspProThrLysGlyGlyLysLysAlaArg	2600	QY	2941	AlaTyrLysSerArgAlaArgAlaValArgAlaSerLeuLeuIleSerArgGlyGlyArgAla	2960
DB	8081	ATGGCCAAAATGAGGTGTTCTGCTGGACCCCAAGGGGGGTAAAGAGCAGCTGCG	8140	DB	9161	GGTGGAAAGCGGGGACGTCAGTCAAGGGCTCCCTCATCTCCCGTGGGGGAGAGCG	9220
				QY	2961	AlaValCysGlyArgTyrLeuPheAsnTyrAlaValLysThrLysLeuLysLeuThrPro	2980

Db 9221 GCGGTTTGGCGCGATATCTCTTCAACTGGCGGTGAAGACCAAGCTCAAACTCACTCCA 9280
 Qy 2981 LeuProGluAlaArgLeuLeuAspLeuSerSerTrpPheThrValGlyAlaGlyGly 3000
 Db 9281 TTCCGCGAAGCGGCGCTCTGGGATTTATCAGCTGGTTCACTGCGCGCGCGCGGGG 9340
 Qy 3001 AspIleThrHisSerValSerArgAlaAaArgProArgLeuLeuPheGlyLeuLeu 3020
 Db 9341 GACATTATCACAGCGTGTGCGTGGCGGACCCCGCTTATTACTCTTGGCCTACTCCTA 9400
 Qy 3021 LeuPheValGlyValGlyLeuPheLeuLeuProAlaArg 3033
 Db 9401 CTTTTGTAGGGTAGGCGCTTTCTTACTTCCCGCTCGG 9439

RESULT 8
 112861
 LOCUS 9589 bp DNA linear PAT 26-JUL-1995
 DEFINITION Sequence 2 from patent US 5428145.
 ACCESSION 112861
 VERSION 112861.1 GI:910242
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 9589)
 AUTHORS Okamoto, H. and Nakamura, T.
 TITLE Non-A, non-B, hepatitis virus genome, polynucleotides,
 polypeptides, antigen, antibody and detection systems
 JOURNAL Patent: US 5428145-A 2 27-JUN-1995;
 FEATURES Location/Qualifiers
 source 1..9589
 BASE COUNT 1968 a 2820 c 2635 g 2166 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 9589
 Score: 15824.00 Matches: 2966
 Percent Similarity: 98.78% Conservative: 30
 Best Local Similarity: 97.79% Mismatches: 37
 Query Match: 98.16% Indels: 0
 DB: 6 Gaps: 0

US-09-980-559-2 (1-3033) x 112861 (1-9589)
 Qy 1 MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgProGln 20
 Db 341 ATGAGCACAAATCTAAACCTCAAGAGAAACCAAGAAACCAACCAACCGTCCGACAA 400
 Qy 21 AspValLysPheProGlyGlyGlyGlnIleValGlyValTyrLeuLeuProArgArg 40
 Db 401 GACGTTAAGTTTCCGCGCGCGCGCAGATCGTTGGCGGAGTATCTTTGTCGCGCAGG 460
 Qy 41 GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly 60
 Db 461 GCGCCAGAGTTGGTGTGGCGCGCAGAGAGACTTCGAGCGGTCCCGACGACGCGGA 520
 Qy 61 ArgArgGlnProLysProLysAspArgSerThrGlyLysSerTrpGlyLysProGly 80
 Db 521 AGCGCGCAGCCCATCCCTAAGAGTCGCGCTCCACTGGCAATCTCTGGGAAACACAGGA 580
 Qy 81 TyrProTrpProLeuTyrGlyAsnGluGlyLeuGlyTrpAlaGlyTrpLeuLeuSerPro 100
 Db 581 TACCCCTGGCCCTATACGGGAATGAGGAGCTCGGCTGGCAGGATGGCTCTGTCCTCC 640
 Qy 101 ArgGlySerArgProSerTrpGlyProAsnAspProArgHisArgSerArgAsnValGly 120
 Db 641 CGAGGTTCCCGTCTCTCTGGGGCCCAATGACCCCGGCAATAGTCCCGCAACGCTGGT 700
 Qy 121 LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProValVal 140
 Db 701 AAGGTCATCGATACCCATAACGTCGGCTTTGCGGACCTCTATGGGGGTATACCTCTGCTGTA 760

Qy 141 GlyValProLeuGlyGlyValAlaAaArgAlaLeuAlaHisGlyValArgValLeuGluAsp 160
 Db 761 GCGCGCGCGCTCGCGCGGTGCCAGAGCTCTCGCGCATGGCGTGAGAGTCTCTGGAGGAC 820
 Qy 161 GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla 180
 Db 821 GCGGTTAATTTTGCACAGGAACTTACCCGGTGTCTCTCTTTTATCTCTTTGCTGGCC 880
 Qy 181 LeuLeuSerCysIleThrThrProValSerAlaAlaGluValLysAsnIleSerThrGly 200
 Db 881 CTGCTGTCTGTCAATCACCAACCGGTCTCCGTGCCGAAGTGAAGACATCATGATCCGGC 940
 Qy 201 TyrMetValThrAsnAspCysThrAsnAspSerIleThrTrpGlnLeuGlnAlaVal 220
 Db 941 TACATGGTGACCAACGACTGCACCAATGATAGCATTACCTGGCACTCCAGGCTGCTGC 1000
 Qy 221 LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTrpIle 240
 Db 1001 CTCACGTCCTCCCGGTGCGTCCGTCGAGAAAGTGGGGAATACATCTCGGTCTGGATA 1060
 Qy 241 ProValSerProAsnValAlaValGlnArgProGlyAlaLeuThrGlnGlyLeuArgThr 260
 Db 1061 CCGGTCTCACCAATGTGGCGGTGCAGCAGCCGCGCTCACGAGGCTTACGAGCG 1120
 Qy 261 HisIleAspMetValValMetSerAlaThrLeuCysSerAlaLeuTyrValGlyAspLeu 280
 Db 1121 CACATTGACATGGTGTGATGTCCGCGACGCTCTGCTCGCTCTTACGTGGGGACCTC 1180
 Qy 281 CysGlyGlyValMetLeuAlaAlaGlnMetPheIleValSerProGlnHisIleTrpPhe 300
 Db 1181 TCGGTGGGTGTATGCTTGCAGCCAGATGTTCAATGTCTCGCACAGCACCACCTGGTT 1240
 Qy 301 ValGlnAspCysAsnCysSerIleTyrProGlyThrIleThrGlyHisArgMetAlaTrp 320
 Db 1241 GTGCAAGACTGCAATTTGCTCCATCTACCTGGTACCATCCTGAGCAGCAGCTGGGTGG 1300
 Qy 321 AspMetMetMetAsnTrpSerProThrAlaThrMetIleLeuAlaTyrAlaMetArgVal 340
 Db 1301 GACATGATGATGAACCTGTCGCGCGCGCTACCATGATCTCTGGCGTACGGATGCGGTC 1360
 Qy 341 ProGluValIleIleAspIleIleSerGlyAlaHisTrpGlyValMetPheGlyLeuAla 360
 Db 1361 CCGAGGTCAATCATAGACATCATTTGGCGGGGTCTATTGGGCGCTCATGTTGGCTTAGCC 1420
 Qy 361 TyrPheSerMetGlnGlyAlaTrpAlaLysValValIleLeuLeuAlaAlaGly 380
 Db 1421 TACTTCTCTATGCGAGGCGTGGGCAAGTGTGTGTCATCTTTTGTCTGGCGCGCGG 1480
 Qy 381 ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr 400
 Db 1481 GTGACGCGCAACCACTACCGTTGGGGTCTTACCGCGCATAAACGCCAGGACCTCAC 1540
 Qy 401 SerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTrp 420
 Db 1541 GGCATGTTCTCCTCTGTCGAGCGCAAAATCCAGCTCATCAACCAATGCGAGTTGG 1600
 Qy 421 HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer 440
 Db 1601 CACATCAACGCGCACCCCTGAACTGTAATGACTTTTGCACACCGGCTCTCTCGCGTCA 1660
 Qy 441 LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg 460
 Db 1661 CTGTTCTTACACCCACAGCTTCACTGTGAGGATGTCCGGAACGATGTCTCGCTGCGC 1720
 Qy 461 SerIleGluAlaPheArgValGlyTrpGlyAlaLeuGlnTyrGluAspAsnValThrAsn 480
 Db 1721 AGTATCGAGGCTTTTGGGTGGGATGGGCGCTTACAAATATGAGGACAAATGACCAAT 1780
 Qy 481 ProGluAspMetArgProTyrCysTrpHisTyrProArgGlnCysGlyValValSer 500
 Db 1781 CCAGAGGATATGAGCGGTATTGCTGGCACTACCCACCAAGACGAGTGTGTGTAGTCTCC 1840

Qy	501	AlaLysThrValCysGlyProValTyrCysPheThrProSerProValValGlyThr	520	Db	2921	GTCCAGAGTGGGACACCTATGTCAGGTGCGCGTGGCGGATCATATGGGCC	2980
Db	1841	GCAGAGTCTGTGTGTGGCCAGTACTGTCTTACCCCCAGCCAGTAGTGGGTACG	1900	Qy	881	ValAlaIlePheTyrProGlyValValPheAspIleThrLysTrpLeuAlaValLeu	900
Qy	521	ThrAspArgLeuGlyAlaProThrTyrThrTyrGlyGluAsnGluThrAspValPheLeu	540	Db	2981	GTCCGATATATCTGCCCCGGGTGTGGTGTGTGATTAACCAAGTGGGTCTTTGGCGGTGCTT	3040
Db	1901	ACCGATAGACTTGGAGCGCCCACTTACACGTGGGGGAGAAATGAGACAGATGCTTCCCTA	1960	Qy	901	GlyProAlaTyrLeuLeuLysGlyAlaLeuThrArgValProTyrPheValArgAlaHis	920
Qy	541	LeuAsnSerThrArgProProLeuGlySerTyrPheGlyCysThrTyrMetAsnSerSer	560	Db	3041	GGGCTGTCTTATCTCTAAAGGTGCTTTGACGCGGTGTCGGTACTTCTCGTCAAGGCTCAC	3100
Db	1961	TTGACAGCACTCGACACCGCAGGGGTGATGGTTCGGCTGCACGTGGATCACTCCACT	2020	Qy	921	AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyGlyArgTyrValGlnMet	940
Qy	561	GlyTyrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer	580	Db	3101	GCTTACTACTAAGGATGTGACCATGTGTAAGGCATCTCGCGGGGTAGTACGTCCAGATG	3160
Db	2021	GGCTACACCAAGACTTGGCGCGCACCACTCGCGCATTAGAGCTGACTTCAATGCCAGC	2080	Qy	941	AlaLeuLeuAlaLeuGlyArgTyrThrGlyTyrIleTyrAspHisLeuThrProMet	960
Qy	581	ThrAspLeuLeuCysProThrAspCysPheArgIleHisProAspThrThrTyrLeuLys	600	Db	3161	GTGCTACTAGCCCTTGGCAGGTGGACTGGCATTATCTATGACCACCTCCACCCCTATG	3220
Db	2081	ATGAGACTTGTGTGCCCCACCGACTGTTTGTAGGAAGCATCTCGATACCATCATCAAA	2140	Qy	961	SerAspTyrAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIleIlePhe	980
Qy	601	CysGlySerGlyProTyrLeuThrProArgCysLeuIleAspTyrProTyrArgLeuTyr	620	Db	3221	TCGGATTGGCTGTCTAATGCGCTGCGGACTTGGCGGTGCGGTGAGGCTTATCATCTTC	3280
Db	2141	TGTGGCTCTGGGCGCTGGCTCAGCCAGGTGCTGATCGACTACCCCTACAGGCTCTGG	2200	Qy	981	SerProMetGluLysLysValIleValTyrGlyAlaGluThrAlaAlaCysGlyAspIle	1000
Qy	621	HisTyrProCysThrValAsnTyrThrIlePheLysIleArgMetTyrValGlyGlyVal	640	Db	3281	AGTCCGATGCGAGAAAAAGTCATCGTCTGGGAGCGGAGACAGCTGCTTGGCGGATATC	3340
Db	2201	CATTACCCCTGCACAGTTAACTATACCATCTTCAAAATAGGATGTATGTGGGGGGTC	2260	Qy	1001	LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp	1020
Qy	641	GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp	660	Db	3341	TTACACGGAGTCTCCCGTGTCCCGGACTTGGCGGGAGGTCTCTTGGCCAGCTGAT	3400
Db	2261	GAGCACAGGCTCAGCGTGGTGCATTTTCACTCGTGGGATCGTTGCACTTTGGAGGAC	2320	Qy	1021	GlyTyrThrSerLysGlyTyrSerLeuLeuAlaProIleThrAlaTyrAlaGlnGlnThr	1040
Qy	661	ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGluTyrAlaIleLeuPro	680	Db	3401	GGCTATACCTCCAAAGGGTGGAGTCTTCTCGCCCCCATCTACTGCTTATGCCAGCAGACA	3460
Db	2321	AGAGACAGAAGTCACTGTCTCTTGTCTGCTGCTCCACACGAGGTGGGCCATTTTACCT	2380	Qy	1041	ArgGlyLeuLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGluGlnAla	1060
Qy	681	CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisLeuHisGlnAsnIle	700	Db	3461	CGCGGCTTTTGGGACCATAGTGTGTGAGCATGACGGGCGGCAAGACAGAACAGGCC	3520
Db	2381	TGCACTTACTCGGACCTGCCCGCTTGTGACTGGTCTTCTCCACCTCCACCAAAACATC	2440	Qy	1061	GlyGluIleGluValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly	1080
Qy	701	ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrLysTyrIleValArgTyr	720	Db	3521	GGGGAGATTGAGTCTGCTCCACGGTCACTAGTCTCTCTCGAAACACCATCTCGGGG	3580
Db	2441	GTGAGCGTGCATTTGATGATGCTTATGCTTCTCTCAAAATACATCGTCCGATGG	2500	Qy	1081	ValLeuTyrThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro	1100
Qy	721	GluTyrValIleLeuLeuPheLeuLeuAlaAspAlaArgValCysAlaCysLeuTyr	740	Db	3581	GTCTTATGAGCTGTCTACCATGAGCTGCGAACAGACTCTAGCCGGCTCACGGGGTCCG	3640
Db	2501	GAGTGGGTAGTACTTATTCTCTGCTTGTAGCGAGCGCCAGGTGTGGCGCTCTATGG	2560	Qy	1101	ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTyrProSerProGly	1120
Qy	741	MetLeuIleLeuLeuGlyGlnAlaGluAlaAlaLeuGluLysLeuValIleLeuHisAla	760	Db	3641	GTCCACACAGATGTACTCCAGTGTGAGGGGGACTTAGTGGGGTGGCCAGCCCCCGGG	3700
Db	2561	ATGCTCATCTTGTGGCCAGGCGGAGCAGCACTAGAGAAGTTGGTGTGCTTGCAGCT	2620	Qy	1121	ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArgAsn	1140
Qy	761	AlaSerAlaAlaSerCysAsnGlyPheLeuTyrPheValIlePhePheValAlaAlaTyr	780	Db	3701	ACCAATCTTTGAGCGCGTGCACGTGTGGAGCGGTGCAGCTATACCTGTGTGCGGAAAC	3760
Db	2621	GCGAGCGCAGTACTGCAATGGCTTCTCTATATCTTGTCTATCTTTTGTGGTGTCTGG	2680	Qy	1141	AlaAspValIleProAlaArgArgArgGlyAspLysArgGlyAlaLeuLeuSerProArg	1160
Qy	781	TyrIleLysGlyArgValValProLeuAlaThrTyrSerLeuThrGlyLeuTyrSerPhe	800	Db	3761	GCTGATGTCTATCCCGCTCGAAGACGCGGGGACAGCGGAGGAGCGGTACTCTCCCCGAGA	3820
Db	2681	TACATCAGGGTGGGTGAGTCCCTTGGCTTATTTCTCTCTGCTGCTGCTGCTGCTGCT	2740	Qy	1161	ProLeuSerThrLeuLysGlySerSerGlyGlyProValLeuCysProArgGlyHisAla	1180
Qy	801	SerLeuLeuLeuAlaLeuProGlnGlnAlaTyrAlaTyrAspAlaSerValHisGly	820	Db	3821	CCTCTTTCCACTTGAAGGGGTCTCTCGGGGGGCGCGGTGCTCTGCCCCCAGAGGCCACGT	3880
Db	2741	GGCCTACTGCTCTAGCATTTGCCCAACAGGCTTATGCTTATGACGATCTGTACATGGT	2800	Qy	1181	ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle	1200
Qy	821	GlnIleGlyAlaAlaLeuValMetIleThrLeuPheThrThrLeuThrProGlyTyrLys	840	Db	3881	GTCCGGGTCTTCCGGGACCGCTGTGCTCCCGGGCGGTGGCCAAAGTCCATAGATTTTATC	3940
Db	2801	CAGATAGGACAGCTCTGTGGTACTGATCACTCTCTTTTCACTCACTCCCGGGGTATAAG	2860	Qy	1201	ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro	1220
Qy	841	ThrLeuLeuSerArgPheLeuTyrTyrLeuCysTyrLeuLeuThrLeuGlyGluAlaMet	860	Db	3941	CCGCTTGACACACTTGACATCTGCTACTCGGTCCCCCACCCTTTAGTGACAAACAGCACCA	4000
Db	2861	ACCCCTCTAGCCGGTCTGTGGTGTGTGCTATCTTCTGACCCCTGGCGGAGCTATG	2920	Qy	1221	ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLys	1240
Qy	861	ValGlnGluTyrAlaProProMetGlnValArgGlyGlyArgAspGlyIleIleTyrAla	880				

4001 CCTGCTGTGCCCCAACTTATCAGGTGGGTACTTACATGCCCCCGACTGTGTAGTGAAG 4060
 1241 SerThrIysValProValAlaIleValAlaGlnGlyTyrIysValLeuValLeuAsnPro 1260
 4061 AGCACCAGTCCCTGTCGGTATGCGCTCAGGGGTACAAAGTGTAGTGTATATCCC 4120
 1261 SerValAlaAlaIleThrLeuGlyPheGlyAlaTyrLeuSerIysAlaHisGlyIleAsnPro 1280
 4121 TCGGTGGTCCACCCCTGGGGTGGGGGTGCTTGTTCGAGGACATGTCATCAATCCC 4180
 1281 AsnIleArgThrGlyValArgThrValThrGlyAlaProIleThrIleThrIleThrIle 1300
 4181 AACATTAGGACTGGGGTCAGGACTGTGACGACCGGGCGCCATCAGTACTCCACATAT 4240
 1301 GlyIysPheLeuAlaAspGlyGlyCysAlaGlyValAlaTyrAspIleIleIleCysAsp 1320
 4241 GGCAGATTCCTCGCCGATGGGGGTGGCGAGGGCGGCCCTATGACATCATATGCGAT 4300
 1321 GluCysHisAlaValAspSerThrIleLeuGlyIleGlyThrValLeuAspGlnAla 1340
 4301 GAATGCCATGCCGTGGACTTACACCATTCCTCGGCATCGAAGACATCTCGATCAAGCA 4360
 1341 GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProGlySerValThr 1360
 4361 GAGACAGCGGGGTACGGCTAACTGTACTGGCTACGGCTACGGCCCGCGGGTCAGTGACA 4420
 1361 ThrProHisProAsnIleGluGluValAlaLeuGlyGlnGlyGluIleProPheTyr 1380
 4421 ACCCCCCACCCCAACATAGAGGAGGTGGCCCTCGGGCAGGAGGTGAGATCCCTTCTAT 4480
 1381 GlyArgAlaIleProLeuSerTyrIleIleGlyArgHisLeuIlePheCysHisSer 1400
 4481 GGGAGGGCGATTCCTCTGTATACATCAAGGAGGAGACATTTGATCTTCTGCCACTCA 4540
 1401 LysIysIysCysAspGluLeuAlaAlaAlaLeuArgGlyMetGlyLeuAsnSerValAla 1420
 4541 AAGAAAAGTGTGACGAGCTCGCGCGGCCCTTCGGGGTATGGGCTTGAACGCGAGTGCCA 4600
 1421 TyrTyrArgGlyLeuAspValSerValIleProThrGlnGlyAspValValValAla 1440
 4601 TACTACAGAGGGGTGGAGCTCTCGTAAATACCAACTCAGGAGACGTAGTGGTGGTGGCC 4660
 1441 ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnVal 1460
 4661 ACCGACGCCCTCATGACGGGGTTTACTGGAGACTTTCGCTGGATCGATCGACGTA 4720
 1461 AlaValThrGlnValValAspPheSerLeuAspProThrPheThrIleThrGlnIle 1480
 4721 GCGGTCACTCAAGTGTAGACTTCAGCTTGGACCCACATTCACATTAACACACACAGACT 4780
 1481 ValProGlnAspAlaValSerArgSerGlnArgArgGlyThrGlyArgGlyArgLeu 1500
 4781 GTCCCTCAAGACGCTGTCTCAGTAGCAGCGCGGGCGCCACGCGGCGAGGAGAGACTG 4840
 1501 GlyIleTyrArgTyrValSerThrGlyValArgAlaSerGlyMetPheAspSerValVal 1520
 4841 GGTATTTATAGTATGTTTCCACTGGTGGCGAGCCCTCAGGAATGTTTGACAGTGTAGT 4900
 1521 LeuCysGluCysTyrAspAlaGlyAlaAlaTyrPyrGluLeuThrProSerGluThrThr 1540
 4901 CTCTGCGAGTCTACGATGACGAGGGCGCATGTTGCTGTGTCGCAAGACCATCTTGAG 4960
 1541 ValArgLeuArgAlaTyrPheAsnThrProGlyLeuProValCysGlnAspHisLeuGlu 1560
 4961 GTCAGGCTCAGAGCATATTTCAACACACCTGTTGCTGTGTCGCAAGACCATCTTGAG 5020
 1561 PheTyrGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThr 1580
 5021 TTTTGGAGGAGGAGTTTTCACCGGCTCACACATAGATGCCACTTCTTCTCCCAACA 5080
 1581 LysGlnSerGlyGluAsnPheAlaTyrLeuThrAlaTyrGlnAlaThrValCysAlaArg 1600
 5081 AAGCAATCGGGGAAAAATTTCCGCATCTTAACAGCCTTACAGGCTACAGTGTGGCTAGG 5140

1601 AlaIysAlaProProSerTyrAspValMetTyrLysCysLeuThrArgLeuLysPro 1620
 5141 GCCAAGACCCCCCGCTCTGGGAGCGTCAITGTGAAGTGTGTGACTCAAGCCC 5200
 1621 ThrLeuValGlyProThrProLeuTyrArgLeuGlySerValThrAsnGluValThr 1640
 5201 ACATCGTGGGCCCCACACTCTCTGTACCGCTTGGGCTCTGTGTACCAACGAGGTCAAC 5260
 1641 LeuThrHisProValThrIysTyrIleAlaThrCysMetGlnAlaAspLeuGluValMet 1660
 5261 CTCACGATCTCTGTGAGAAATACATCGCCACTGTCATCAAGCCGACCTTGAAGTCAATG 5320
 1661 ThrSerThrTyrValLeuAlaGlyGlyValLeuAlaAlaValAlaAlaIleTyrCysLeuAla 1680
 5321 ACCAGCAGCTGGGTCTTAGCTGGGGGGTCTTGGGGCGCGTGGCCGCTACTGCTGGCG 5380
 1681 ThrGlyCysValCysIleIleGlyArgLeuHisIleAsnGlnArgAlaValValAlaPro 1700
 5381 ACCGGGTGTGTTCATCATCGCGCTTGCACGCTTAACGAGCGCGCTCGTTGCAACG 5440
 1701 AspLysGluValLeuTyrGluAlaPheAspGluMetGluGluCysAlaSerArgAlaAla 1720
 5441 GACAAGGAGGCTCTCTATGAGGCTTTGTATGAGATGGAGAAATGTCCTCTAGAGCGCT 5500
 1721 LeuIleGluGluGlyGlnArgIleAlaGluMetLeuLysSerIysIleGlnGlyLeuLeu 1740
 5501 CTCTATTGAAGAGGGGCGAGCGGATAGCCGAGATGCTGAAGTCCAAGATCCAAGGCTTATTG 5560
 1741 GlnGlnAlaSerIysGlnAlaGlnAspIleGlnProThrValGlnAlaSerTyrProLys 1760
 5561 CAGCAAGCTTCCAAACAAGCTCAAGACATACACCGCTGTGAGGCTTCTTGGCCCAAG 5620
 1761 ValGluGlnPheTyrAlaIysHisMetTyrAsnPheIleSerGlyIleGlnTyrLeuAla 1780
 5621 GTAGAGCAATTCGGGCCAAACACATGTGAATCTTCATCAGCGGCAATCAATACCTCGCA 5680
 1781 GlyLeuSerThrLeuProGlyAsnProAlaValAlaSerMetMetAlaPheSerAlaAla 1800
 5681 GGNATATCAACACTGCCGAGAACCTCTGTGTAGCTTCCATGATGGCATTCAGTGGCGCC 5740
 1801 LeuThrSerProLeuSerThrThrThrIleLeuLeuAsnIleLeuGlyGlyTyrLeu 1820
 5741 CTCAACGATCCGTTGTCAACTAGCACCATCTCTCTCAACATTTTGGGGGCTGGCTA 5800
 1821 AlaSerGlnIleAlaProProAlaGlyAlaThrGlyPheValIleSerGlyLeuValGly 1840
 5801 GCATCCCAATTCGCGCTCCCGGGGGCTTACCGGCTTCTGTCTAGTGGCTTGGTGGG 5860
 1841 AlaAlaValGlySerIleGlyLeuGlyLysValLeuValAspIleLeuAlaGlyTyrGly 1860
 5861 GCTGCGGTAGGACATAGCTTGGTGAAGTGTCTGTGGACATCTTGGCAGGATGTGT 5920
 1861 AlaGlyIleSerGlyAlaLeuValAlaPheLysIleMetSerGlyGluLysProSerMet 1880
 5921 GCGGCAATTCGGGGGCTCTCTCGTCGATTCAGATCATGTCTGGCGAGAGCCCTCCATG 5980
 1881 GluAspValValAsnLeuLeuProGlyIleLeuSerProGlyValAlaLeuValGlyVal 1900
 5981 GAGGATGTGTCAACTGTCTGCTGCTGAAATCTCTCTCGGGGTGCTTGGTGGGAGTC 6040
 1901 IleCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTyrMet 1920
 6041 ATCTGCGGGCCATCTCGCCCGACAGCTGGAGCGGGGAGCGCTGTCTCAATGATG 6100
 1921 AsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaProThrHisTyrValThr 1940
 6101 AATAGGCTCATTCCTTGTCTTCCAGAGGAAACACAGCTCGCCCGCCCACTAGTACG 6160
 1941 GluSerAspAlaSerGlnArgValThrGlnLeuGlySerLeuThrIleThrSerLeu 1960
 6161 GAGTCGATGCGTCGAGCGGTGTGACCACTACTTGGCTCTCTTACCAATACCAAGCCTG 6220

QY	1961	LeuArgLeuHisAsnTrpIleThrGluAspCysProIleProCysGlySerTrp	1980	Db	7301	AAAAACCCGACGCTCCCGACAGGAGACGCGGAGTGGTCTCGAGTGAGAGCTCCATA	7360
Db	6221	CTCAGGAGACTCCACAACCTGATTAAGAGCTGCCCATCCATCCAGCGGCTCGTG	6280	QY	2341	GlyAspAlaLeuGlnLeuAlaIleLysSerPheGlyGlnProProSerGlyAsp	2360
QY	1981	LeuArgAspValTrpAspTrpValCysThrIleLeuThrAspPheLysAsnTrpLeuThr	2000	Db	7361	GCAGATGGCCCTAACAGCTGGCCATCAAGTCTTTGGCAGCCCCCCCCAGCGCGAT	7420
Db	6281	CTCCGGATGTGGGATGGGTGGACCATCTAACAGACTTTAAAAACTGGCTGACC	6340	QY	2361	SerGlyLeuSerThrGlyAlaGlyAlaAspSerGlySerGlnThrProProAspGlu	2380
QY	2001	SerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTrpLys	2020	Db	7421	TCAGGCTTTTCCACGGGGCGGACGCGAGTCCGCGAGTCGGACGCCCCCGCATGAG	7480
Db	6341	TCCAAATGTTCCCAAGATGGCTGGTCTCCCTCTTATCTTGTCAAAGGGGTACAG	6400	QY	2381	LeuAlaLeuSerGluThrGlySerIleSerSerMetProProLeuGluGlyLeuGly	2400
QY	2021	GlyValTrpAlaGlyThrGlyIleMetThrThrArgCysProCysGlyAlaAsnIleSer	2040	Db	7481	TTGGCCCTTTTCGAGACAGGTTCCATCTCTCCATGCCCTCTCTCGAGGGGAGCTGGA	7540
Db	6401	GGCGTGGGCTGGCACTGGTATCATGACACACACGCGTCTCTTCGCGCGCAATATCTCT	6460	QY	2401	AspProAspLeuGluProGluGlnValGluProGlnProProGlnGlyValAla	2420
QY	2041	GlyAsnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTrp	2060	Db	7541	GATCCAGACTTGAGGCTGAGCAGGTAGAGCTTCAACCTCCCCCGGGGGGGTGGTA	7600
Db	6461	GGCAATGTCCGCTGGGCTCATGAGAAATACGGGGGCCMAAACCTGCATGAATATCTGG	6520	QY	2421	AlaProGlySerAspSerGlySerTrpSerThrCysSerGluGluAspAspSerVal	2440
QY	2061	GlnGlyThrPheProIleAsnCysTrpThrGluGlyGlnCysValProLysProAlaPro	2080	Db	7601	ACCCCGGCTCAGGCTCGGGGTCTTGGTCTACTTGTCTCGAGGAGGACGACTCCGTCGTG	7660
Db	6521	CAGGGACCTTTCCCATCAATGTTACACGGAGGGCCAGTGGCTGCCGAAACCCGACCA	6580	QY	2441	CysCysSerMetSerTrpSerTrpThrGlyAlaLeuIleThrProCysSerProGluGlu	2460
QY	2081	AsnPheLysValAlaIleTrpArgValAlaAlaSerGluTrpAlaGluValThrGlnHis	2100	Db	7661	TGCTGTCTCATGCTCATCTCTCGACCGGGGCTCTAATAACTCTCTGTAGCCCCGAAGAG	7720
Db	6581	AACTTTAAGATCGCCATCTCGAGGGTGGCGCTCAGAGTACGCGGAGGTGACGACGAC	6640	QY	2461	GluLysLeuProIleAsnProLeuSerAsnSerLeuLeuArgTrpHisAsnLysValTyr	2480
QY	2101	GlySerTrpHisTrpIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeu	2120	Db	7721	GAAGAAGTTGCCAATTTGGCCCCCTTGAGCAACTCCCTGTTGCGATATCACAAACAGGTGATC	7780
Db	6641	GGGTCTATACCACATACATAACAGGACTTACCACATGAATACTTGAAGTTCTCTGCCAACA	6700	QY	2481	CysThrThrThrLysSerAlaSerLeuArgAlaLysLysValThrPheAspArgMetGln	2500
QY	2121	ProSerProGluPhePheSerTrpValAspGlyValGlnIleHisArgPheAlaProThr	2140	Db	7781	TGTACCATCATCAAGAGCGCCTCTTAAGGGCTTAAAGAGTAACCTTTTGATAGGATGCAA	7840
Db	6701	CTTCTCCAGATTCCTTCTCGGTGGAGCGAGTGCAGATCCATAGGTTTGCCCCCATTA	6760	QY	2501	ValLeuAspSerTrpTrpAspSerValLeuLysAspIleLysLeuAlaLysLysVal	2520
QY	2141	ProLysProPhePheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal	2160	Db	7841	CGGCTCGAGCTCATTTATGACTCAGTCTTGAAGGACATTAAGCTAGCGGCTCCAAAGTTC	7900
Db	6761	CCGAAGCCGTTTTTTCGGGATGAGTCTCGTCTGCGTGGGCTTAATTCATTTGTCGTC	6820	QY	2521	ThrAlaArgLeuLeuThrMetGluAlaCysGlnLeuThrProProHisSerAlaArg	2540
QY	2161	GlySerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThr	2180	Db	7901	ACCGAAGGCTTCTCCTTACGAGGAGGCTCCAGTTAACTCCACCCCACTTCGAAGA	7960
Db	6821	GGGTCTCAGCTCCCTTGCAGTCTGAACTCGACACAGACGATATTGACGTCATGCTAACA	6880	QY	2541	SerLysTrpGlyPheGlyAlaLysGluValArgSerLeuSerGlyArgAlaValAsnHis	2560
QY	2181	AspProSerHisIleThrAlaGluThrAlaAlaArgLeuAlaArgGlySerProPro	2200	Db	7961	TCCNAGTATGGTTTGGGCTAAGGAGGTCGCGAGCTTCTCCGGAGAGCCGTTAACCAC	8020
Db	6881	GACCATCCCATATACCGCGAGACTGACGCGCGGCTTTGGACGCGGGTCAACCCCGG	6940	QY	2561	IleLysSerValTrpLysAspLeuLeuGluAspSerGluThrProIleProThrThrIle	2580
QY	2201	SerGluAlaSerSerAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThr	2220	Db	8021	ATCAAGTCGCTGGGAAGGACCTCTCTGGAAGACACACAAACCAATTCCTACCAACCATC	8080
Db	6941	TCCGAGGCAAGCTCCTCAGCGAGCCAGCTATCGGCACCATCGCTGCGAGCCACTGCACC	7000	QY	2581	MetAlaLysAsnGluValPheCysValAspProThrLysGlyGlyLysLysAlaAlaArg	2600
QY	2221	ThrHisGlyLysAlaTrpAspValAspMetValAspAlaAsnLeuPheMetGlyGlyAsp	2240	Db	8081	ATGGCCAAATAATGAGTGTCTTCGCTGGACCCCAACAGGGGGGTAAAGAAGCAGCTCGC	8140
Db	7001	ACCCACGGCAAGGCTATGATGTGGACATGGTGGATGCCAACCTGTTTCATGGGGGGCGAT	7060	QY	2601	LeuIleValTrpProAspLeuGlyValArgValCysGluLysMetAlaLeuTrpAspIle	2620
QY	2241	ValThrArgIleGluSerGlySerLysValValLeuAspSerLeuAspProMetVal	2260	Db	8141	CTTATCGTTTACCTGACCTCGGCTCAGGGTCTGGAGAAATATGGCCCTTTATGATATC	8200
Db	7061	GTGACCGGATAGAGTCTGAGTCAAAGTGGTCTCTTGAGCTCTCTCGACCCCAATGGTC	7120	QY	2621	ThrGlnLysLeuProGlnAlaValMetGlyAlaSerTrpGlyPheGlnTrpSerProAla	2640
QY	2261	GluGluArgSerAspLeuGluProSerIleProSerGluTrpMetLeuProLysLysArg	2280	Db	8201	ACACAAAAGCTTCTCAGCGGCTGATGGGGCTTCTTATGGATTCCAGTACTCCCCCGCT	8260
Db	7121	GAAGAAAGGAGCGACCTTGAGCCTTCGATACCATCGGAATATATGCTCCCAAGAGAGA	7180	QY	2641	GlnArgValGluPheLeuLeuLysAlaTrpAlaGluLysLysAspProMetGlyPheSer	2660
QY	2281	PheProProAlaLeuProAlaTrpAlaArgProAspTrpAsnProProLeuValGluSer	2300	Db	8261	CAGCGGTGAGTCTTCTTGAAGGATGCGGGGAAAAGAAAGACCTTATGGGTTTTTCG	8320
Db	7181	TTCCACACGCTTACCGGCTGGGACGCGCTGATTAACAACCCACCGCTTGTGGAATCG	7240	QY	2661	TyrAspThrArgCysPheAspSerThrValThrGluArgAspIleArgThrGluGluSer	2680
QY	2301	TrpLysArgProAspTrpGlnProAlaThrValAlaGlyCysAlaLeuProProArg	2320	Db	8321	TATGATACCCGATGCTTTGACTCAACCGTCACTGAGAGAGACATCAGGACTGAGGAGTCC	8380
Db	7241	TGAAGAGGCGCATTAACCAACCGGCACTGTTGGGGGCTCGCTCTCCCCCCCCCTAAG	7300	QY	2681	IleTrpArgAlaCysSerLeuProGluGluAlaHisThrAlaIleHisSerLeuThrGlu	2700
QY	2321	LysThrProThrProProArgArgArgThrValGlyLeuSerGluAspSerIle	2340				

8381 ATATATCGGCTTGTCTTCCCGGAGGAGCCACACTGCCATACACTCACTGACTGAG 8440
 2701 ArgLeuTyrValGlyGlyProMetPheAsnSerIleGlyGlnThrCysGlyTyrArgAg 2720
 8441 AGACTTTTACGTGGAGGCGCCATGTTCAACAGAGGCGCCAGACCTCGCGGTACAGCGGT 8500
 2721 CysArgAlaSerGlyValLeuThrThrSerMetGlyAsnThrIleThrCysTyrValLys 2740
 8501 TGCGCGCCAGCGGGTGTCTTACCATTAGCATGGGAACCATCATGCTATGTGAA 8560
 2741 AlaLeuAlaCysLysAlaAlaGlyIleAlaProThrMetLeuValCysGlyAsp 2760
 8561 GCCTTAGCGCTGTAAAGCTGAGGATATTCGCGCCCAATGTGTGTATCGCGCAT 8620
 2761 AspLeuValIleSerGlySerGlnGlyThrGluAspGluArgLeuAla 2780
 8621 GACTTGGTGTCTATCTAGAGAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 8680
 2781 PheThrGluAlaMetThrArgTyrSerAlaProGlyAspProProArgProGluTyr 2800
 8681 TTACGAGGAGCTATGACGAGTATTCGCGCTTCTGTCGAGGAGGAGGAGGAGGAGTAT 8740
 2801 AspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaLeuGlyProGlnGly 2820
 8741 GACCTGAGCTGATACATCTGCTCTCAATGTGTCTGCTGCTGCTGCTGCTGCTGCTG 8800
 2821 ArgArgArgTyrLeuThrArgAspProThrThrProIleAlaArgAlaAlaTyrGlu 2840
 8801 CGCGCAGATACCTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8860
 2841 ThrValArgHisSerProValAsnSerThrLeuGlyAsnIleIleGlnTyrAlaProThr 2860
 8861 ACAGTTAGACATCTCCCTGTCAATTCATGCTAGGAGGAGGAGGAGGAGGAGGAGGAG 8920
 2861 IleTrpAlaArgMetValLeuMetThrHisPhePheSerIleLeuMetAlaGlnAspThr 2880
 8921 ATATGGCTCGCATGCTGCTGATGACACATCTTCTTCTTCTTCTTCTTCTTCTTCTT 8980
 2881 LeuAspGlnAsnLeuAsnPheGluMetTyrGlyAlaValTyrSerValSerProLeuAsp 2900
 8981 CTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 9040
 2901 LeuProAlaIleLeuGluArgLeuHisGlyLeuAspAlaPheSerLeuHisThrTyrThr 2920
 9041 CTCCAGCCATATTTGAAGGTACAGGCTTACAGGCTTCTTCTTCTTCTTCTTCTTCTT 9100
 2921 ProHisGluLeuThrArgValAlaSerAlaLeuArgLysLeuGlyAlaProProLeuArg 2940
 9101 CCCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 9160
 2941 AlaTrpLysSerArgAlaArgAlaValArgAlaSerLeuIleSerArgGlyGlyArgAla 2960
 9161 CGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 9220
 2961 AlaValCysGlyArgTyrLeuPheAsnTrpAlaValLysThrLysLeuLysLeuThrPro 2980
 9221 GCCGTTTCCGCGCCATATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 9280
 2981 LeuProGluAlaArgLeuAspLeuSerSerTrpPheThrValGlyAlaGlyGlyGly 3000
 9281 TTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 9340
 3001 AspIleTyrHisSerValSerArgAlaArgProArgLeuLeuLeuPheGlyLeuLeuLeu 3020
 9341 GACATTTATATACAGCGGTGCGGTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 9400
 3021 LeuPheValGlyValGlyLeuPheLeuLeuProAlaArg 3033
 9401 CTTTGTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 9439

RESULT 9
 AF169005
 LOCUS AF169005 9700 bp RNA linear VRL 06-SEP-2001

DEFINITION Hepatitis C virus isolate NDM59 polyprotein gene, complete cds.
 ACCESSION AF169005
 VERSION AF169005.1 GI:6707285
 KEYWORDS Hepatitis C virus
 SOURCE Hepatitis C virus
 ORGANISM Hepacivirus.
 REFERENCE 1 (bases 1 to 9700)
 AUTHORS Kurihara, C., Ishiyama, N., Nishiyama, Y., Fukushi, S., Kageyama, T., Katayama, K., and Miura, S.
 TITLE Molecular characterization of hepatitis C virus genotype 2a from the entire sequences of four isolates
 JOURNAL J. Med. Virol. 64 (4), 466-475 (2001)
 MEDLINE 21361470
 PUBMED 11468731
 REFERENCE 2 (bases 1 to 9700)
 AUTHORS Katayama, K. and Kurihara, C.
 TITLE Direct Submission
 JOURNAL Submitted (14-JUL-1999) Basic Research Division, Biomedical Laboratories, Inc., 1361-1 Matoba, Kawagoe, Saitama 350-1101, Japan
 FEATURES
 Location/Qualifiers
 1..9700
 /organism="Hepatitis C virus"
 /mol_type="genomic RNA"
 /isolate="NDM59"
 /db_xref="taxon:11103"
 /note="genotype: 2a"
 341..9442
 /codon_start=1
 /product="polyprotein"
 /protein_id="AAF25613.1"
 /db_xref="GI:6707286"
 /translation="MSTNPKPORKTKENTSRPQDVKPGGQIVGGVLLPRRGRL
 GYARTRKTSRQSRQPIKDRRSTGSKGKPGYPMPLYNBEGELGAGWLLSPRG
 SPSPGPTDRHRSNRVNGKVIDTLTCGFDALMGYTPVVGAGLGGVARALAHGVRL
 GYNTATGPNLPGCSFIFLLALSCIPFVSAVQVKNISDSYMTNDNSDITWQLOA
 AVLHGGVPCPKMGNISCMIPVSNVAVRPGALTQGLRAHIDVNMVNSATLCSALY
 VDLGCGVMAAQMFTVSPQHFVQECNSIVPGALTGRHAWDMVMNSATPMTAIL
 AYAMRVEVIIDIISGAWGVMLGAYFSKQAWAKVIVLILTAGDARTSIAGSV
 AHATSLAGLFTSGAKQNIQLINTNGSHINTALNCNDSLNTGTFIASLTYFRFNS
 GCPRLSAGCIQAFRIQWGTURIEDNVNTPEDMRPYCWHYPPKCGIVSARSVCGPV
 YCFPSVVGVTDRIGVPTWGENETDVFILNTPRPGSGMFCGFWNNSGTFTKTC
 GAPPRIADFNASMDLLCPTCFRKHDPATYIKCGSPWLPRLCLVDYPTFLMHPFC
 TNYTIFKRMVGVGVEHRLTAACRTGDPGNCLEDRDRSOLSPILLHSTWAILPCS
 YSDPALSTGLLHLHQNIVDQMYGSLPALTXYIVRWEVLLFLLADARVACALM
 MLILQAEALAEKLVHLHAASACNGFLYFVIFVAAWYIKRAVPLAAYSLTGLM
 PFCLLLALPQAYAYADSVHGTGVALILILITFLTPTGKILLRSLSWMLCYLTL
 GEAMTQEWYPPMARGGRDGIITAAATMFCGVVFDITKLLMLGAPYLLRDLATRPV
 YFVRAHALRLCTMVRHLAGRYVQMALALGRWTGTYTYDHLSPMSDWAASGLRDLA
 VAVEPIIRSPMEKVIWGAETAACDVLHGLPVSARLGREILLGADGVTSGKWRLL
 APITAYAQTRGLGAIIVSMTGRDKEQGEIQLVSTVTSQFLGTSISGLVTVYHG
 ANKTLASGRGPTQMSAEGDLVGPSPGKTSLEPCTCGAVALDYLIVTRNADVIPA
 RRRGKRRLSPRLSTLKGSGGVLCPRGHVGIFRAAVCSRGVAKSIDFIPVEA
 LDVTRGTPFSDNTPPNQRTQVTHLHPTGSKSTKVPVAYAAQGVKVLVNPVS
 AATLGRVYLSKAHGINPRTQVTVTTGESITTYTGKFLADGQACAGADYIICD
 ECHADVATILGVTVLDOAETAGVRLTVLATATPGSVTTPHNEEVALGQEGRI
 FVGRALPSYIIGKRHLIFCHSKKDELSAALRSNGLNAVAYRYRGLDVSIVTQGDV
 VVATDALMTGTYGDFSDVNDVAVTVQVDFSLDPTFTITQTVQDDAVSRQRGR
 TGRGLIGIYVYVTSGERASGDFSVLCEYDAGAAWELTSETTVLRAYRNPFTGL
 PQCDHLEFWEAVFTGLTIDAHFSLQSGENFAYLVAQYATCARAKAPPNDV
 MKCLTRLKTLVGLPTLRGLPNTVLTTHPTKYTATQADLEVMVTSVTLVAG
 GYLAVAACVLAGTGVSIIGRLHVNQAVVAPDKEVLYEAFDEMECASRAALIIEQ
 RAEMLSKIQGLLOQASQADQIPAVOASWPKVQFPAKHWNVNLSIGLYLAGLST
 LPENAVASMAFSAALTSPLSTLILNLGGWLASLAPPAGATGFWVSGLVGNA
 VSGIIGKGLVDILAGYAGISALVAFKIMSGEKPSMEDVINLLPGIISPGALVGV
 ICALILRRHVGPCEGAVQWNRILAFASRGNHVAPHTYTESDASQVTLQGLSLIT
 SCAALLRRHMTEDCPICAGSLREVDWVCTILTFKNWLSKLPKMPGLFISQ
 KGVGVWAGTGINTRPCGANTSGNVRLSGMRIITGPKTCMNTQGTFFPNCYTEGQ
 VKPAPNFKTAIRVAASVAYEATVHGSYITGLTDSIKVPCQPSPEFSWGVG
 QHRFAPTTPKPFDRVSPFCVGLNSFVVGSLPCDPEPDADVLTLMLTDSHTAENA
 ARRLARGSPSEASASLSAPSLRATCTTHGKAYDVMVDANLFWGGDVSRSEY
 KVVVLDLSLDPMAERSDLFSPSETYMLPRNRPFPALPAWARPPPLVPEWKRPDY

QPTVAGCALPPPKKTTTPPPRRRTVGLNENTIGDALQOLAIAKFGQPLSDGSLGS
 TGDAADSRTSPDESALSSETGISISSMPLEGEQDPOLEPEQVLSLPPQEEYVP
 GDSGSWSTCSDEDDSVVSCMSYSWTGALITPCSPDEEKLPIPLNSLLRYHNKVY
 CTTKGSASLRKAVTDRMOVLDAHVDKIKLAASKVSARLLTLEBACOLTPPHS
 ARSKYGFGEKVESLSGRAVNHKISKVKOLLDSQTPITPMKNEVEFPVPAKGGK
 KAARLIVPDLGVRCBKALYDVDTKLOAVMGASYGFQYSPAQRVEFLKAWAEKK
 DPMGFSYDTRCFSTVTERDIRTEESYIRACSLPREAHTAIHSLTERLYVGGPMLNSK
 GQTCGYRRCSRASGLVTTMGNTITCYIKALAAACAAGIVAPTMLVCGDDLVISESQ
 TEEDERNLRAFTAMTRYGAPDPPRPEYDLELITSCSNVSVAXXGPOGRRRYLYLR
 DPTPLSRAMETVRHSPVNSLGNIIQVAPTIVWRMLWTHFFSILMAQDITDONLN
 FEMYGVSYSPLDLPAI LERHGLDAFSLHTYSPELFRVASLKLKCAPLRAWKS
 RARAVRASLISRGRAAICGRYLFENWAVTKLPLPEARLLDLSSWFTVAGCGGDI
 YHSVSRAPKLLLSLLLSLVGVGLFLPAR"

BASE COUNT 1948 a 2845 c 2712 g 2190 t 5 others
 ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 9700
 Score: 15426.00 Matches: 2874
 Percent Similarity: 97.13% Conservative: 72
 Best Local Similarity: 94.76% Mismatches: 87
 Query Match: 95.69% Indels: 0
 DB: 14 Gaps: 0

US-09-980-559-2 (1-3033) x AF169005 (1-9700)

Qy 1 MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGln 20
 Db 341 ATGAGCACAAATCTTAACCTCAAGANAACCAAGANAACACGACCGCTGCCCCACA 400
 Qy 21 AspValLysPheProGlyGlyGlnIleValGlyValLysValLysLeuLeuProArg 40
 Db 401 GAGCTTAAGTTCCGGGTGGCGCCAGATCGTTGGCGGAGTATACTTGTTGGCGCGCAGG 460
 Qy 41 GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly 60
 Db 461 GGCCCCAGGTTGGGTGTGGCGCGGACGAGGAAGATTCGGAGCGGTCCCCAGCCAGCTGGA 520
 Qy 61 ArgArgGlnProLysAspArgArgSerThrGlyLysSerThrGlyLysProGly 80
 Db 521 AGAGCCGAGCCCATCCCAAGATCGGCGCTCCACTGGCMAAGTCTTGGGGNAACACAGGA 580
 Qy 81 TyrProTrpProLeuTyrGlyAsnGluGlyLeuGlyTyrAlaGlyTyrLeuLeuSerPro 100
 Db 581 TACCCCTGGCCCTGTATGGGAATGAGGACTCGGCTGGCGAGGGTGGCTCTGTCTCCCC 640
 Qy 101 ArgGlySerArgProSerTrpGlyProAsnAspProArgHisArgSerArgAsnValGly 120
 Db 641 CGAGGTTCCCGTCTCTCGGGGCCCACTAGACCCCGGCATAGGTGCGGCACACGTGGGT 700
 Qy 121 LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProValVal 140
 Db 701 AAGGTATCGACACCTTACGTGCGGCTTTGGCGACTCATGGGGTACATCCCGCTCGTA 760
 Qy 141 GlyAlaProLeuGlyGlyValAlaArgAlaLeuAlaHisGlyValArgValLeuGluAsp 160
 Db 761 GGGCGCCCGCTTGGTGGCGTGGCAGAGCTCTCGCGCAGCGGTGAGAGTCTCTGGAGGAC 820
 Qy 161 GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla 180
 Db 821 GGGGTTAAATATGCAACAGGAACTTACCTGGTGTCTCTCTTTCTATCTTTCTGGTGGCC 880
 Qy 181 LeuLeuSerCysIleThrThrProValSerAlaAlaGluValLysAsnIleSerThrGly 200
 Db 881 TTAAGTCTGCAATACCGTTCCGGTTTACGGCTTCAAGTGAAGAACATCAGCGACAGC 940
 Qy 201 TyrMetValThrAsnAspCysThrAsnAspSerIleThrTrpGlnLeuGlnAlaVal 220
 Db 941 TATATGCTGACTAATGACTCTCTAATGACAGCATTTACCTGGCAGCTTCAGGCGCGGTC 1000
 Qy 221 LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTrpIle 240
 Db 1001 CTCACGTCCTCCCGGTGGTGGTGTGAGAAATGGAAACATATACGCGTCTCGATA 1060

Qy 241 ProValSerProAsnValAlaValGlnArgProGlyAlaLeuThrGlnGlyLeuArgThr 260
 Db 1061 CCGGTTTACCAAAACGTGGCTGTGCGAGCGCTTGGCGCCCTCAGCGCAGGCTTGGCGGG 1120
 Qy 261 HisIleAspMetValValMetSerAlaThrLeuCysSerAlaLeuTyrValGlyAspLeu 280
 Db 1121 CACATCGACATGGT 1180
 Qy 281 CysGlyGlyValMetLeuAlaAlaGlnMetPheIleValSerProGlnHisIleTrpPhe 300
 Db 1181 TGTGGCGGGGT 1240
 Qy 301 ValGlnAspCysAsnCysSerIleTyrProGlyThrIleThrGlyHisArgMetAlaTrp 320
 Db 1241 GTGAGGAATGCAATTTGCTTCCATCTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1300
 Qy 321 AspMetMetMetAsnTrpSerProThrAlaThrMetIleLeuAlaTyrAlaMetArgVal 340
 Db 1301 GACATGATGATGAACCTGGTGGCGCCAGCCACCATGATCTTGGCGGTACGCGATGCGGTT 1360
 Qy 341 ProGluValIleIleAspIleIleSerGlyAlaHisTrpGlyValMetPheGlyLeuAla 360
 Db 1361 CTTGAGTCTCATGACATCATTTAGCGGGCTCAGTGGGGCTCATGTTGGCTTGGCC 1420
 Qy 361 TyrPheSerMetGlnGlyAlaTrpAlaLysValValIleLeuLeuLeuAlaGly 380
 Db 1421 TACTTCTCTATGCAAGGACATGGCGAAAGTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1480
 Qy 381 ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr 400
 Db 1481 GTGAGCGGCATATCGCGCAGCATTTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1540
 Qy 401 SerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTrp 420
 Db 1541 GGCTCTTCACTTCTGGCGCTAAGCAGAAACATCCAGCTCATTAACACCAATGGCAGTGG 1600
 Qy 421 HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer 440
 Db 1601 CACATCAACCGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1660
 Qy 441 LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg 460
 Db 1661 CTGTTCTACACCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 1720
 Qy 461 SerIleGluAlaPheArgValGlyTyrGlyAlaLeuGlnTyrGluAsnThrValThrAsn 480
 Db 1721 GGTATCCAGGCTTCCGATAGGGTGGGCGACCTTGGGATACGAGGATACGTCACCAAC 1780
 Qy 481 ProGluAspMetArgProTyrCysTrpHisTyrProProArgGlnCysGlyValValSer 500
 Db 1781 CCAGAAGATATGAGACCATCTGCTGGCACTATCCCAACCAACAGTGTGGCATAGTCTCC 1840
 Qy 501 AlaLysThrValCysGlyProValTyrCysPheThrProSerProValValValGlyThr 520
 Db 1841 GCGAGTCTGTGTGTGGCCAGTACTGTCTTCACTTCACTTCACTTCACTTCACTTCACTT 1900
 Qy 521 ThrAspAspLeuGlyAlaProThrTyrThrTyrGlyGluAsnGluThrAspValPheLeu 540
 Db 1901 ACCGACAGGCTCGGAGTGTCTTACAGTGGGAGAGAAATGAGACGAGTGTCTTCATA 1960
 Qy 541 LeuAsnSerThrArgProProLeuGlySerTrpPheGlyCysThrTrpMetAsnSerSer 560
 Db 1961 TTGAACAGACCCCGACCCCGGGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2020
 Qy 561 GlyTyrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer 580
 Db 2021 GGTTCACCAAGACCTGTGGCGCAGCGCTGTGGCGCATTTAGAGCTGACTTCAATGCCAGC 2080
 Qy 581 ThrAspLeuLeuCysProThrAspCysPheArgLysHisProAspThrThrTyrLeuLys 600
 Db 2081 ATGACGCTGTGTGTGGCCCGGACTGCTTTAGGAAGCATCTGTGTGTGTGTGTGTGTGT 2140

Qy 601 CysGlySerGlyProThrLeuThrProArgCysLeuIleAspTyrProTyrArgLeuTrp 620
 Db 2141 TGTGGTCTGGGCGCTGGCTCAGCCAGGTGCTGGTGGCTACCTTACAGGCTCTGG 2200
 Qy 621 HisTyrProCysThrValAsnTyrThrIlePheLysIleArgMetTyrValGlyGlyVal 640
 Db 2201 CATTACCCCTGCACATCACTATACCATCTTCAAGATAAGGATGATGTGGGGGGGTT 2260
 Qy 641 GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp 660
 Db 2261 GAGCACAGGCTCAGCGCGCATCACTTCACTCGTGGGATCCTTGCACCTTGGAGGAC 2320
 Qy 661 ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGluTrpAlaIleLeuPro 680
 Db 2321 AGGACAGAGCCACACTGCTCTCCTTGTGTCATCTCACACGGAATGGGCCATTTTGCCT 2380
 Qy 681 CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisLeuHisGlnAsnIle 700
 Db 2381 TGCTCTTACTCAGACTTGGCCCGCTTGTGCTGACTGGCTTCTTCTATCTCCACCAAAATATC 2440
 Qy 701 ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrIleLysTyrIleValArgTrp 720
 Db 2441 GTGACCTACAGTACATGATGCTGCTGCTACCTGCTCTCAAAAATACATCGTTCGGTGG 2500
 Qy 721 GluTrpValIleLeuLeuPheLeuLeuAlaAspAlaArgValCysAlaCysLeuTrp 740
 Db 2501 GAGTGGGTAGTACTTCTTATTTCTGCTCTTGGCGGATGCCAGGCTGCGCTTGTGGTGG 2560
 Qy 741 MetLeuIleLeuLeuGlyGlnAlaGluAlaLeuGluLysLeuValIleLeuHisAla 760
 Db 2561 ATGCTCATCTTGTGGCGCAGCGCGGAGCAGCAGTACAGAAAGCTGGTCTTGCATGCC 2620
 Qy 761 AlaSerAlaIleAspCysAsnGlyPheLeuTyrPheValIlePhePheValAlaIleTrp 780
 Db 2621 GCGAGCGCAGCCAGCTGATAAGCTTCTATATTTGTGCTATCTTTTCTGGCTGCTGG 2680
 Qy 781 TyrIleLysGlyArgValValProLeuAlaThrTyrSerLeuThrGlyLeuTrpSerPhe 800
 Db 2681 TACATCAAGGTCGGCGGCTCCCTTGGCGCTTATTTCCCTCACTGCTTGTGGCCCTTT 2740
 Qy 801 SerLeuLeuLeuAlaLeuProGlnAlaTyrAlaTyrAspAlaSerValHisGly 820
 Db 2741 TGCCTGTGTCTCTAGCACTGCCCAACAGCTTATGATATGATGATCGTGCAGCA 2800
 Qy 821 GlnIleGlyAlaAlaLeuLeuValMetIleThrLeuPheThrLeuThrProGlyTyrLys 840
 Db 2801 CAGACAGCGTGGCTTTGTGATACCTATACCTCTTCACTTACCCCGGGGTATAAG 2860
 Qy 841 ThrLeuLeuSerArgPheLeuTrpTrpLeuCysTyrLeuLeuLeuThrLeuGlyGluAlaMet 860
 Db 2861 ATCTCTTCTAGCCGGTCTCTGTGGTGTGTGCTATCTCTGACTCTGGGGGAGGCCATG 2920
 Qy 861 ValGlnGluTrpAlaProMetGlnValArgGlyGlyArgAspGlyIleIleTrpAla 880
 Db 2921 ATCCAGAGTGGGTACCACTCCAGCATGCGCGGTGTGCGCGATGATCATATGGCC 2980
 Qy 881 ValAlaIlePheTyrProGlyValValPheAspIleThrLysTrpLeuLeuAlaValLeu 900
 Db 2981 GCCACCATGTTCTCCCGGGGTAGTATTTGACATAACCAAGTGGCTCTTGGCGATGCTT 3040
 Qy 901 GlyProAlaTyrLeuLeuLysGlyValAlaLeuThrArgValProTyrPheValArgAlaHis 920
 Db 3041 GGGCCCCGTATCTCTCGGAGATGCTCTGACACGCTGCGTATTTTGGTCAGAGCTCAC 3100
 Qy 921 AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyGlyArgTyrValGlnMet 940
 Db 3101 GCTCTGTGAGTGTGCTCATGTCAGGATCTCGCGGGGTAGTACGTCCAGTCATG 3160
 Qy 941 AlaLeuLeuAlaLeuGlyArgTyrThrGlyThrTyrIleTyrAspHisLeuThrProMet 960
 Db 3161 GCGCTATTAGCCCTTGTGTAGTGGACCGGCTTACATCTATGACCACTCTCCCTATG 3220
 Qy 961 SerAspTrpAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIleIlePhe 980

Db 3221 TCGGATGGCTGTAGCGGCTGGGGAGCTTGGCGGTGGAGCCCATCATCTTC 3280
 Qy 981 SerProMetGluLysIleValIleValTrpGlyAlaGluThrAlaAlaCysGlyAspIle 1000
 Db 3281 AGTCCGATGGAGAAAGTCACTGCTGGGGAGGGAGAGCGGTGCATGCGGGGACGTC 3340
 Qy 1001 LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp 1020
 Db 3341 TTGCACGAGACTTCCGTGTCGCCGACTTGGCCGGGAGATCTCTTGGCCAGCTGAT 3400
 Qy 1021 GlyTyrThrSerLysGlyTrpSerLeuLeuAlaProIleThrAlaTyrAlaGlnGlnThr 1040
 Db 3401 GGCTACACCTCTAAGGGTGGAGGCTTCTTGGCCCCATCACTGCTTACGCCACGACACA 3460
 Qy 1041 ArgGlyLeuLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGluGlnAla 1060
 Db 3461 CGAGGTCTCTTGGCGCTATAGTGTGAGCATGACGGGCGTGACAAAGACAGAACAGGCC 3520
 Qy 1061 GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly 1080
 Db 3521 GGGGAAATCCAAAGTCTTATCCACAGTCACTCACTCTTCTCGGAACATCCATTTGGGG 3580
 Qy 1081 ValLeuTrpThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro 1100
 Db 3581 GTCTTATGACTGTTTATCATGAGCTGGCAATAAACTCTAGCCGCTCGCGGGCCCG 3640
 Qy 1101 ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTrpProSerProGly 1120
 Db 3641 GTCACGACATGCTACTCGAGCGCGAGGGGAGCTTGTAGGGTGGCCAGCCCTCTGGG 3700
 Qy 1121 ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArgAsn 1140
 Db 3701 ACCAAATCTTTGAGCCGTGCACATGTGGAGGGTGCAGCTGCTGTGTCACGCGGAC 3760
 Qy 1141 AlaAspValIleProAlaArgArgArgGlyAspLysArgGlyAlaLeuLeuSerProArg 1160
 Db 3761 GCTGATGTATCCCGCCGAGAGAGCGGGGCAAGCGGGGAGCGCTACTCTCCCGGAGA 3820
 Qy 1161 ProLeuSerThrLeuLysGlySerSerGlyGlyProValLeuCysProArgGlyHisAla 1180
 Db 3821 CCTCTTTCAGCTTGAAGGGGTCTTCGGGGGAGCCGCTCTTGGCTTGGGGCCAGCT 3880
 Qy 1181 ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle 1200
 Db 3881 GTCCGATTTTCCGGCAGCTGTGCTCTCGGGGGGTGGTAAAGTCCATAGATTTCATC 3940
 Qy 1201 ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro 1220
 Db 3941 CCGGTTGAGCGCTCGAGCTCGTCACTGCTCCCACTTCTAGTGACAAACAGCACACCA 4000
 Qy 1221 ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLys 1240
 Db 4001 CCGGCGCTGCTCAGACCTATCAGGTCCGGTACTTGCATGCCCAACCGCGAGTGAAG 4060
 Qy 1241 SerThrLysValProValAlaTyrAlaGlnGlyTyrLysValLeuValLeuAsnPro 1260
 Db 4061 AGTACCAAGTCTCTCGGTACCGCTCAGGGGTACAAAGTACTAGTGTATATCCC 4120
 Qy 1261 SerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerLysAlaHisGlyIleLeuAsnPro 1280
 Db 4121 TCGGTGCTGCCACCTCGGGTTCGGGGCTACTTGTCCAAGGCACATGTCATCAATCCT 4180
 Qy 1281 AsnIleArgThrGlyValArgThrValThrThrGlyAlaProIleThrTyrSerThrTyr 1300
 Db 4181 AACATTAGGACTGGAGTCAGAACCCGACCGGGGAGTCCCATCATCATCTCCAGTAC 4240
 Qy 1301 GlyLysPheLeuAlaAspGlyGlyCysAlaGlyValAlaTyrAspIleIleLeuCysAsp 1320
 Db 4241 GGGAAATTTCTCTGATGGGGCTGTGACGGGGGGGCTATGACATCATCATATGCGAT 4300
 Qy 1321 GluCysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGlnAla 1340

Db 4301 GAATGCCACGCTGTGGATGCTACACCAATCTCGGCGTCGGAACGGTTCTTGACCAAGCA 4360
 Qy 1341 GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProGlySerValThr 1360
 Db 4361 GAGACAGCGGGGTACAGACTAACCGTCTGSCACAGGCTACGCCCCCTGGTGGTGACA 4420
 Qy 1361 ThrProHisProAsnIleGluGluValAlaLeuGlyGlnGluGlyGluIleProPheThr 1380
 Db 4421 ACCCCCCCATCCCAATATAGAGGAGGTAGCCCTCGGCGAGGAGGCGAGATCCCTCTCTAT 4480
 Qy 1381 GlyArgAlaIleProLeuSerThrIleLeuGlyGlyArgHisLeuIlePheCysHisSer 1400
 Db 4481 GGGAGGCGCATTCCTCTCTTACATTAAGGAGGAGGAGGACCTGATTTTCTGCCACTCA 4540
 Qy 1401 LysLysLysCysAspGluLeuAlaAlaLeuArgGlyMetGlyLeuAsnSerValAla 1420
 Db 4541 AAGAAAAGGTGACGAGCTCTCGGCGCTCTTCGGAGCATGGCTTGAATGCTGTGGCA 4600
 Qy 1421 TyrTyrArgGlyLeuAspValSerValIleProThrGlnGlyAspValValValAla 1440
 Db 4601 TACTACAGAGGGTGGACGCTCCGTAATACCAACTCAGGAGACGCTAGTGGTGTGCGC 4660
 Qy 1441 ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAsnSerValIleAspCysAsnVal 1460
 Db 4661 ACCGACGCCCTCATGACGGGTATCTGGAGACTTCGACTCCGTGATCGACTGCAACGTA 4720
 Qy 1461 AlaValThrGlnValValAspPheSerLeuAspProThrPheThrIleThrGlnIle 1480
 Db 4721 GCGGTCACTCAAGTCTGGACTTCAGCTTAGACCCCACTTTTCACTATACCAACACAGACT 4780
 Qy 1481 ValProGlnAspAlaValSerArgSerGlnArgGlyArgThrGlyArgGlyArgLeu 1500
 Db 4781 GTCCCGCAAGACGCTGTCTCACGTAGCAGCGCGGGCGGCACAGGTAGAGAGACGCTG 4840
 Qy 1501 GlyIleTyrArgTyrValSerThrGlyGluArgAlaSerGlyMetPheAspSerValVal 1520
 Db 4841 GGCATTTATAGGTATGTTTCCATCTGGTGGAGCGCTTCAGGAATGTTTGACGTAGTG 4900
 Qy 1521 LeuCysGluCysTyrAspAlaGlyAlaAlaTyrTyrGluLeuThrProSerGluThrThr 1540
 Db 4901 CTCTGTGAGTGTACGACGCGAGGGCTGTCTGTATGAGCTCACCATCGGAGACTACC 4960
 Qy 1541 ValArgLeuArgAlaTyrPheAsnThrProGlyLeuProValCysGlnAspHisLeuGlu 1560
 Db 4961 GTCAGGCTTAGCGGTATTTCAACAGCCCTGTGCTGTGTCAAGATCATCTTGAG 5020
 Qy 1561 PheTyrGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThr 1580
 Db 5021 TTTTGGAGGAGCTCTTTACTGGCCCTCACACATAGATGCCCACTTCTCTTTCCCAACA 5080
 Qy 1581 LysGlnSerGlyGluAsnPheAlaTyrLeuThrAlaTyrGlnAlaThrValCysAlaArg 1600
 Db 5081 AAGCATCGGGGAGGAATTCGATATCTAGTAGCCCTATCAGGCTACAGTGTGCGCCAGA 5140
 Qy 1601 AlaLysAlaProProSerThrProAspValMetTyrLysCysLeuThrArgLeuLysPro 1620
 Db 5141 GCCAAAGCCCTCCCTCGCTCGGAGCTCATGTGAAGTGTCTTGACTCAAGGCC 5200
 Qy 1621 ThrLeuValGlyProThrProLeuLeuTyrArgLeuGlySerValThrAsnGluValThr 1640
 Db 5201 ACACCTGGGGCCCTACACCTCTCTCTGACCGTTTGGGGCTGTGTATACCAACGAGGTCACT 5260
 Qy 1641 LeuThrHisProValThrLysThrIleAlaThrCysMetGlnAlaAspLeuGluValMet 1660
 Db 5261 CTCACATCTCTGTGAGAAATACATCGCCACATGATGCAAGCTGACCTTAGGTCATG 5320
 Qy 1661 ThrSerThrTrpValLeuAlaGlyValLeuAlaAlaValAlaAlaTyrCysLeuAla 1680
 Db 5321 ACCAGCAGCTGGGTCTTAGCTGGGGAGTCTCGCAGCGCTGCTCGGTATTGCTTGGCG 5380
 Qy 1681 ThrGlyCysValCysIleIleGlyArgLeuHisIleAsnGlnArgAlaValAlaPro 1700
 Db 5381 ACCGGGTGCGTTTCCATCATTTGGCCGCTTGACAGTCAACAGCAGCGCGCTTGACCG 5440

Qy 1701 AspLysGluValLeuTyrGluAlaPheAspGluMetGluGluCysAlaSerArgAlaAla 1720
 Db 5441 GACAAAGAGTCTCTCTATGAGCGGTTTCGAGAGATGGAAGATGTCCTCTAGAGCGGCT 5500
 Qy 1721 LeuIleGluGluGlyGlnArgIleAlaGluMetLeuLysSerLysIleGlnGlyLeuLeu 1740
 Db 5501 CTCATTGAAGAGGGGAGCGAATAGCTGAGATGCTGAAGTCCAAGATCCAAGGCTTATTG 5560
 Qy 1741 GlnGlnAlaSerLysGlnAlaGlnAspIleGlnProThrValGlnAlaSerTrpProLys 1760
 Db 5561 CAGCAAGCTCTTAACAGGCCCCAGGATATACAAACCGCTGTGAGGCTCTATGGCCCAAG 5620
 Qy 1761 ValGluGlnPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeuAla 1780
 Db 5621 GTGGAGCAATTCGTGGSCCAACACACATGCGAATTCATTAGTGGCATCCAAATATCTCGCA 5680
 Qy 1781 GlyLeuSerThrLeuProGlyAsnProAlaValAlaSerMetMetAlaPheSerAlaAla 1800
 Db 5681 GGACTGTCAACACTGCGCAGGGAACCTGCTGTGGTTCATGTATGGCATTCAGTGGCGCC 5740
 Qy 1801 LeuThrSerProLeuSerThrThrThrIleLeuLeuAsnIleLeuGlyTyrLeu 1820
 Db 5741 CTCACAGTCCGTGTGTGACTAGCACCTATCTCTCAACATATTTGGAGGCTGGCTG 5800
 Qy 1821 AlaserGlnIleAlaProProAlaGlyAlaThrGlyPheValValSerGlyLeuValGly 1840
 Db 5801 GCGTCCCAATCGCGCGCGCGCGGCGCACTGGCTTCTGTCAGTGGCTAGTGGGG 5860
 Qy 1841 AlaAlaValGlySerIleGlyLeuGlyLysValLeuValAspIleLeuAlaGlyTyrGly 1860
 Db 5861 GCTGCGGTAGGAGCATAGGCTGGTAAAGTGTGGTGGATATCTCGCAGGCTATGGT 5920
 Qy 1861 AlaGlyIleSerGlyValAlaPheLysIleMetSerGlyGlyLysProSerMet 1880
 Db 5921 GCGGCAATTTGGGGGCGCTCTGATAGATTCAGATCATCTGCGGAGAGCGCTCCATG 5980
 Qy 1881 GluAspValValAsnLeuLeuProGlyIleLeuSerProGlyAlaLeuValValGlyVal 1900
 Db 5981 GAGGAGCTCATCAACTGCTGCTGGATCTTGTCTCCGGCGCGCTGTGTAGTGGGAGTC 6040
 Qy 1901 IleCysAlaAlaIleLeuArgHisValGlyProGlyGluGlyAlaValGlnTrpMet 1920
 Db 6041 ATCTGTGCGCATTTCTGCGCGCCACGTGGGGCGGGGAGGCGCGTCCAAATGGATG 6100
 Qy 1921 AsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaProThrHisTyrValThr 1940
 Db 6101 AATAGGCTCATTTGCTTCTCCAGAGGAAACACAGTCCGCCCGCCACCCACTACGTGACG 6160
 Qy 1941 GluSerAspAlaSerGlnArgValThrGlnLeuLeuGlySerLeuThrIleThrSerLeu 1960
 Db 6161 GAGTGGATGCTCGCAGCGCTGACCCACTGCTTGGCTCTCTTACTATTAATAGCTCTG 6220
 Qy 1961 LeuArgArgLeuHisAsnTrpIleThrGluAspCysProIleProCysGlyGlySerTrp 1980
 Db 6221 CTCAGAGACTCCAAATTTGGATTTACTGAAGACTGCCCATCCCATCGCGCGCTCATGG 6280
 Qy 1981 LeuArgAspValTrpAspTrpValCysThrIleLeuThrAspPheLysAsnTrpLeuThr 2000
 Db 6281 CTCGCGAGGTTGGAGCTGGGTTTGCACCATCTTAAACAGACTTTTAAAAAAGCTGGTGACC 6340
 Qy 2001 SerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTyrLys 2020
 Db 6341 TCCAAAGTGTTCCTCCAAAGATGCCGCGCTCTCTTATCTCTCCGCAAAAGGGGTACAGG 6400
 Qy 2021 GlyValTrpAlaGlyThrGlyIleMetThrThrArgCysProCysGlyAlaAsnIleSer 2040
 Db 6401 GCGGTGTGGCGCGCACTGGCATCATGACACACGATGCTCTTGGCGCGCCCAACATTC 6460
 Qy 2041 GlyAsnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTrp 2060
 Db 6461 GGTAATGCTCGGCTCTATGAAATTTACAGGCGCTTAAACCTGCACTGAACACTTGG 6520

2061 GlnGlyThrPheProIleAenCysTyrThrGluGlyGlnCysValProLysProAlaPro 2080
 6521 CAGGGACCTTTCCCATCAATTGTTACACGGAGGCGCAGTGCCTGCGAAACCCCGGCCA 6580
 2081 AsnPhelysValAlaIleTPrArgValAlaAAsrGluTyrAlaGluValThrGlnHis 2100
 6581 AACTTCAAGACCCCATCTCGAGGGTGGCGCTCGAGGTACGCGAGGTGACGAGCAC 6640
 2101 GlySerTyrHisTyrIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeu 2120
 6641 GGGTCATCTCTATATAACAGGACTGACTACTGACAGTTTGAAGTTCCCTGCTCAATTA 6700
 2121 ProSerProGluPhePheSerTPrValAspGlyValGlnIleHisArgPheAlaProThr 2140
 6701 CCTTCTCCAGAGTTTTTCTCGGTGGAGCGAGTGCAGATCCATAGGTTTGCTCCACG 6760
 2141 ProLysProPheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal 2160
 6761 CCAAGACCGTTTTTCCGGGATAGGTCTCGTCTCGGTGGCTCAATTGTTTGTGCTC 6820
 2161 GlySerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThr 2180
 6821 GGGTCTCAGCTTCTCTGTGACCTCGAGCCGCGCAGACGTATTTGACGTCCATGCTAACA 6880
 2181 AspProSerHisIleThrAlaGluThrAlaAlaArgArgLeuAlaArgGlySerProPro 2200
 6881 GATCCGTCCCATATACGCGGGAGGCTGACGCGCGGCTTAGCGGGGATCACCCCGC 6940
 2201 SerGluAlaSerSerAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThr 2220
 6941 TCCGAGCAAGCTCTCGCGAGCCAGCTATCGCACCGCTCGTGGAGCCACTGCACC 7000
 2221 ThrHisGlyLysAlaTyrAspValAspMetValAspAlaAsnLeuPheMetGlyLysAsp 2240
 7001 ACCCAGCGCAAGGCTATGATGTGGAATGCTGATGCGAAGCTGTTTCATGGAGCGCAT 7060
 2241 ValThrArgIleGluSerGlySerLysValValValLeuAspSerLeuAspProMetVal 2260
 7061 GTGACTCGGATAGAGTCTGAGTCCAAAGTGGTCTGCTGAGACTCTCTCGACCCCAATGGCC 7120
 2261 GluGluArgSerAspLeuGluProSerIleProSerGluTyrMetLeuProLysLysArg 2280
 7121 GAAGAAGGAGCGACCTCGAGCTTCAATACCATCGGAGTATATGCTCCCGCAGCAACAGG 7180
 2281 PheProProAlaLeuProAlaTPrAlaArgProAspTyrAsnProProLeuValGluSer 2300
 7181 TTCCCGCAGCTTTACCGGCTGCGGCAAGCGCTGATTAACACCGCGCTGTGGAAATCA 7240
 2301 TrpLysArgProAspTyrGlnProAlaThrValAlaGlyCysAlaLeuProProArg 2320
 7241 TGGAGAGGCCAGATTACCAACCGCTTACTTGTGCGGGCTGTGCTCTCCCGCTCCCAAG 7300
 2321 LysThrProThrProProArgArgArgArgThrValGlyLeuSerGluAspSerIle 2340
 7301 AAGACTCCGAGCGCCCCCAAGGAGAGCTCGACAGTGGGTCTAAACGAGAACACCATTA 7360
 2341 GlyAspAlaLeuGlnGlnLeuAlaIleLysSerPheGlyGlnProProProSerGlyAsp 2360
 7361 GGAGATGCCCTTCAACAGCTGGCCATCAAGGCTTTCGCGCAGCGCCCTCTGAGCGCGCAT 7420
 2361 SerGlyLeuSerThrGlyAlaGlyAlaAlaAspSerGlySerGlnThrProProAspGlu 2380
 7421 TCAGGCTTTTCCAGGGGGGAGCGCGCGAGCTCCGCGAGTCCGAGCGCCCCCGCGATGAG 7480
 2381 LeuAlaLeuSerGluThrGlySerIleSerSerMetProProLeuGluGlyGluLeuGly 2400
 7481 TCGGCTCTTTCGAGACAGGTTCTATCTCTCCATGCCCCCTTCGAGGGGGAGCGCGGG 7540
 2401 AspProAspLeuGluProGluGlnValGluProGlnProProProGlnGlyGlyValAla 2420
 7541 GACCCAGACCTTAGAGCTGAGCAGGTGGAGCTTCCAGCTCCCCCGGAGGAGGAGGTA 7600
 2421 AlaProGlySerAspSerGlySerTPrSerThrCysSerGluGluAspSerValVal 2440

7601 GTTCTCTGGCTCGGACTCGGGTCTTGGTCCACTTGTCTCGAGGAGGATGACTCCGTCGTG 7660
 2441 CysCysSerMetSerTyrSerTPrThrGlyAlaLeuIleThrProCysSerProGluGlu 2460
 7661 TGCTGCTCCATGCTATATTCCTGGACCGGGGCTCTAATAAATCTCTTGTAGCCCCGAAGAG 7720
 2461 GluLysLeuProIleAsnProLeuSerAsnSerLeuLeuArgTyrHisAsnLysValTyr 2480
 7721 GAAAGTTTCCCAATTAACCCCTTGAGTAATCTGCTATTTGCGATATCACAAATAGGTGTAC 7780
 2481 CysThrThrThrLysSerAlaSerLeuArgAlaLysLysValThrPheAspArgMetGln 2500
 7781 TGFACTACATCAAGAGCGCTCTTATAGGGCCCAAGAGGTAACTTTTTCACAGGATGCAA 7840
 2501 ValLeuAspSerTyrThrAspSerValLeuLysAspIleLysLeuAlaLaserLysVal 2520
 7841 GTGCTCGACGCCCTTATGACTCAGTCTTAAAGGACATCAAGCTAGCGGCTTCCCAAGTTC 7900
 2521 ThrAlaArgLeuLeuThrMetGluAlaCysGlnLeuThrProProHisSerAlaArg 2540
 7901 AGCGAAGGCTCTTACCTTGGAGGAGGCTGCCAGTTGACTTCCACCCCATTTCCGCAAGA 7960
 2541 SerLysTyrGlyPheGlyAlaLysGluValArgSerLeuSerGlyArgAlaValAsnHis 2560
 7961 TCCAAAGTATGGTGGGCTTAAGGAGGTCCGAGCTTGTCCGGAGGCGCTCAACCCAC 8020
 2561 IleLysSerValTPrLysAspLeuLeuAspSerGluThrProIleProThrIle 2580
 8021 ATCAAGTCCGTGGAGGACCTCTCGAAGACTCACAACACCAATCTCTACGACATC 8080
 2581 MetAlaLysAsnGluValPheCysValAspProThrLysGlyLysLysAlaAlaArg 2600
 8081 ATGGCCAAAATAGGTGTTCTGCTGGACCTTCCCAAGGGGGGCAAGAAAGCAGCTCGC 8140
 2601 LeuIleValTyrProAspLeuGlyValArgValCysGluLysMetAlaLeuTyrAspIle 2620
 8141 CTATATGCTTACCTGACCTCGCGCTCAGGGTTTGGAGAGATGCGCTTTATGATGTC 8200
 2621 ThrGlnLysLeuProGlnAlaValMetGlyAlaSerTyrGlyPheGlnTyrSerProAla 2640
 8201 ACACAAAGCTTCTCAGCGGTGATGGGGCTTCTACGGCTTCCAGTACTCCCCCGCT 8260
 2641 GlnArgValGluPheLeuLeuLysAlaTPrAlaGluLysLysAspProMetGlyPheSer 2660
 8261 CAACGGGTGGAGTTTCTTAAAGGCTGCGCGGAAAGAACACCTATGGGTTTTTCG 8320
 2661 TyrAspThrArgCysPheAspSerThrValThrGluArgAspIleArgThrGluGluSer 2680
 8321 TATGATACCGGATGCTTGTATTCACCGTCACTGAGAGAGACATTAAGAATGAGGAGTCC 8380
 2681 IleTyrArgAlaCysSerLeuProGluGluAlaHisThrAlaIleHisSerLeuThrGlu 2700
 8381 ATATACCGGCTGCTTCTCGCGGAGGCGCCACACTGCCCATACACTCGCTAACTGAG 8440
 2701 ArgLeuTyrValGlyGlyProMetPheAsnSerLysGlyGlnThrCysGlyTyrArgArg 2720
 8441 AGACTCTAGTGGAGGGGCCATGCTCAACAGCAGAGGGTTCAGACCTTCGCGATACAGCGGT 8500
 2721 CysArgAlaSerGlyValLeuThrThrSerMetGlyAsnThrIleThrCystyrValLys 2740
 8501 TGCGCGCCAGCGGGGTGCTCACCTAGCATGGGGAACACCATCATCATGCTCATATAAAA 8560
 2741 AlaLeuAlaAlaCysLysAlaAlaGlyIleIleAlaProThrMetLeuValCysGlyAsp 2760
 8561 GCCCTAGCGCTTGCAGGCTCGCGGGATAGTTCGCCCTTACCATGTGGTATCGCGGAC 8620
 2761 AspLeuValValIleSerGluSerGlnGlyThrGluGluAspGluArgAsnLeuArgAla 2780
 8621 GACTTGGTTGTTATCTCAGAAAGCCAGGAGCTGAGGAGGACGCGAAACCTTGAGAGCC 8680
 2781 PheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProArgProGluTyr 2800

Copied from 09540843 on 05/19/2004

Db	2381	:	TGCACCTTACTCAGACCTGCCCGCTTGTGACCGGTCTTCTCCACCTCCACCAAAACATC	2440
Qy	701	ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrLysTyrIleValArgTyr	720	
Db	2441	GTGGATGTACATACATATGATGGCTGTACCTGTCTTACAAATATACATCGTTCGATGG	2500	
Qy	721	GluTyrValIleLeuLeuPheLeuLeuLeuAlaAspAlaArgValCysAlaCysLeuTyr	740	
Db	2501	GAGTGGGTGTACTTATTTCTTGTCTCTAGCGGACCGAGGTTTTCGGCTCGCTATGG	2560	
Qy	741	MetLeuIleLeuLeuGlyGlnAlaGluAlaAlaLeuGluLysLeuValIleLeuHisAla	760	
Db	2561	ATGCTCATTTTGTAGCCAGGCGGAAGCAGCACTAGAGAAGTAGTGTCTTTCACGCC	2620	
Qy	761	AlaSerAlaAlaSerCysAsnGlyPheLeuTyrPheValIlePhePheValAlaAlaTyr	780	
Db	2621	CGGAGCGCGGCGAGTCAATGGCTTCTATATTTTGTCTTCTTCTTCTTCTTCTTCTT	2680	
Qy	781	TyrIleLysGlyArgValProLeuAlaThrTyrSerLeuThrGlyLeuTyrSerPhe	800	
Db	2681	TACATCAAGGCGCGGCGTCCCTTGGCTTACCTATTCCCTCACCCTCTATGGTCTTTC	2740	
Qy	801	SerLeuLeuLeuAlaLeuProGlnGlnAlaTyrAlaTyrAspAlaSerValHisGly	820	
Db	2741	TGCTTACTGTCTCGGCAATGGCCCAACAGAGCTTATGCGTATGATACATCTGGCGCAG	2800	
Qy	821	GlnIleGlyAlaAlaLeuLeuValMetIleThrLeuPheThrLeuThrProGlyTyrLys	840	
Db	2801	CAGATAGCGCGAGCTGTCTTAATACTGATTAATCTCTTACACTCACCCCGGCGTATAAG	2860	
Qy	841	ThrLeuLeuSerArgPheLeuTyrTrpLeuCysTyrLeuLeuThrLeuGlyGluAlaMet	860	
Db	2861	ACCTTGTGACCGGTTTTATGGTGGCTGTGCTATCTTCTGACCTTGGGGGAGCCATG	2920	
Qy	861	ValGlnGluTrpAlaProMetGlnValArgGlyGlyArgAspGlyIleIleTrpAla	880	
Db	2921	GTCAGAGGTGGGCACCACTTTGTCAGGTGGCGGTGGCGGTGATGCATCATATGGCC	2980	
Qy	881	ValAlaIlePheTyrProGlyValValPheAspIleThrLysTrpLeuLeuAlaValLeu	900	
Db	2981	CGCCGCATATCTGCGCGGCGGTGGTGTGTGATACCAAGTGGCTCTTAGCGTTGCTT	3040	
Qy	901	GlyProAlaTyrLeuLeuLysGlyAlaLeuThrArgValProTyrPheValArgAlaHis	920	
Db	3041	GGGCTGCTTATCTCTAAGAGATGCTTTGACGGGTGCGGTACTTCTGTCAGAGTCCAC	3100	
Qy	921	AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyGlyArgTyrValGlnMet	940	
Db	3101	GCTCTGCTGAGGATATGCATATGTTAAGGCATCTCGCGGGGGTAGGTACGTCCAGATG	3160	
Qy	941	AlaLeuLeuAlaLeuGlyArgTyrThrGlyTyrIleTyrAspHisLeuThrProMet	960	
Db	3161	CGGCTACTAGCCCTTGGTAGGTGAGCTGGGCACTTACATCTATGACCACTTCAGCCCATG	3220	
Qy	961	SerAspTrpAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIleIlePhe	980	
Db	3221	TCGGATTGGGCTGCCAATGCCCTGCGGACTTGGCGGTCCAGTGGAGCCTATCATCTTC	3280	
Qy	981	SerProMetGluLysLysValIleValTrpGlyAlaGluThrAlaAlaCysGlyAspIle	1000	
Db	3281	AGTCCGATGAGAAGAAAGTCATCGTCTGGGAGCGGAGACGCGTGCATGTGGGAGCATA	3340	
Qy	1001	LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp	1020	
Db	3341	CTACACGAGCTTCCCGTCTGCGCGACTTGGTGGGAGGTCTCTCTGGCCCGAGCTGAT	3400	
Qy	1021	GlyTyrThrSerLysGlyTyrSerLeuLeuAlaProIleThrAlaTyrAlaGlnThr	1040	
Db	3401	GGCTACACCTTCCAGGGGTGGAGTCTTCTCGCCCCCATCTACTCTTACGCCCGCAGACG	3460	
Qy	1041	ArgGlyLeuLeuGlyThrIleValSerMetThrGlyArgAspLysThrGluGlnAla	1060	
Db	4541	AAGAAAAAGTGTGACGAGCTCGCGCGGCGCTTTCGGGGTATGGGCTTGAACGACGATGGCA	4600	
Db	3461	CGGGGTCTTTTGGCGCTATAGTGTGAGCATGACGGGGCGGCGCAAGACAGAACAGGCC	3520	
Qy	1061	GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly	1080	
Db	3521	GGGGAATTCAGGTCTCTGCTACAGTCACTCAGTCTCTCTCGAAACATCCATCGCGGG	3580	
Qy	1081	ValLeuTyrThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro	1100	
Db	3581	GTCTTATGAGCTGTCTATCATGGGGTGGCAACAAGACCTTAGTCTGGCTCAGCGGTCCA	3640	
Qy	1101	ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTyrProSerProGly	1120	
Db	3641	GTCAACGAGTGTACTCCAGTCCGAGGGGGAGCTTGGTGGAGTGGGCGCTCCCTCTCTGG	3700	
Qy	1121	ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArgAsn	1140	
Db	3701	ACTAAGTCTTGGAGCGGTGCAGTGTGGAGCGGTGCAGCTGTACTTGTCTGTCACCGGAC	3760	
Qy	1141	AlaAspValIleProAlaArgArgArgGlyAspLysArgGlyAlaLeuLeuSerProArg	1160	
Db	3761	GCTGAGCTCATCCGGCTCGAAGACGTGGGGAACAAGCGGGAGCACTACTCTCCCGAGA	3820	
Qy	1161	ProLeuSerThrLeuLysGlySerSerGlyGlyProValLeuLysProArgGlyHisAla	1180	
Db	3821	CCACTTTTCAGCTTAAAGGGGTCTCGGGGGGGCGGTCTTTTGGCCCGACGCGCCAGCT	3880	
Qy	1181	ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle	1200	
Db	3881	GTCGGGATCTTTCGGGCGAGCTGTGTCTCTCGGGGGGTGGCTAAGTCCATAGATTTCATC	3940	
Qy	1201	ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro	1220	
Db	3941	CCGTTTGAGACACTCGACATCGTGACGGGTCCCCCACTTTAGTGCACAACAGCACCA	4000	
Qy	1221	ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLys	1240	
Db	4001	CCTGTCTGCCCGCAGACCTATCAGGTGGGTACTTACATGCCCGGCTGTGTAGTGAAG	4060	
Qy	1241	SerThrLysValProValAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnPro	1260	
Db	4061	AGCACCAAGTCTCTGCTGTCTACGCCCGCCAGGGGTATAAAGTGTGTGTCTCAATCCC	4120	
Qy	1261	SerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerLysAlaHisGlyIleAsnPro	1280	
Db	4121	TCGGTGGCGCCACCCCTGGGATTCGGGGGTACCTGTCCAGGACACATGCATCATCCC	4180	
Qy	1281	AsnIleArgThrGlyValArgThrValThrGlyAlaProIleThrTyrSerThrTyr	1300	
Db	4181	AATATCAGGACTGGAGTCAAGGACTGTGACGACCGGAGCGGCATCATCTACTCCATAC	4240	
Qy	1301	GlyLysPheLeuAlaAspGlyGlyCysAlaGlyGlyValaTyrAspIleIleCysAsp	1320	
Db	4241	GGCAATTCCTCGCCGATGGGGGTCTCGGGGGGTCTTATGATATCATCATATCGCAT	4300	
Qy	1321	GluCysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGlnAla	1340	
Db	4301	GAGTGCATTCGGTGGTGTCTACCACTCCCTCGGCATCGGAACAGTCTCTCGACCAAGCA	4360	
Qy	1341	GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProProGlySerValThr	1360	
Db	4361	GAGACAGCGGGGTTCAGGTCAACTGTGTGTGGTCTACGGCCACGCGCCCTCTGGGTGACA	4420	
Qy	1361	ThrProHisProAsnIleGluGluValAlaLeuGlyGlnGluGlyIleProPheTyr	1380	
Db	4421	ACCCCCACCCCAATATAGAGAGGTGGCCCTCGGGCAGAGGGTGGAGATTCTCTCTTCTAC	4480	
Qy	1381	GlyArgAlaIleProLeuSerTyrIleLysGlyGlyArgHisLeuIlePheCysHisSer	1400	
Db	4481	GGGAGGGCGATTCCTCTGTCTTACATCAAGGGGAGGAGACCTGTATCTCTTGCCACTCA	4540	
Qy	1401	LysLysLysCysAspGluLeuAlaAlaLeuArgGlyMetGlyLeuAsnSerValAla	1420	
Db	4541	AAGAAAAAGTGTGACGAGCTCGCGCGGCGCTTTCGGGGTATGGGCTTGAACGACGATGGCA	4600	

QY	1421	TyTyrArgGlyLeuAspValSerValIleProThrGlnGlyAspValValValAla	1440	QY	1781	GlyLeuSerThrLeuProGlyAsnProAlaValAlaSerMetMetAlaPheSerAlaAla	1800
DB	4601	TACTACAGAGGGTTGGACGATCATATACCAACTCAGGGGATGTGGTGTCTGCGC	4660	DB	5681	GGACTGTCAACACTGCCAGGAACCCCGCTGTGGCTTCCATGATGGCATTCAGTGGCGC	5740
QY	1441	ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnVal	1460	QY	1801	LeuThrSerProLeuSerThrSerThrThrIleLeuLeuAsnIleLeuGlyTyrLeu	1820
DB	4661	ACCGAGCCCTCATGACAGGATACACTGGGGACTTTTGACTCTGTGATCGACTCAACGTA	4720	DB	5741	CTCACCAGTCCGTGTCAACTAGCACCACTATCTCTTCAACATTTTGGAGGCTGGCTG	5800
QY	1461	AlaValThrGlnValValAspPheSerLeuAspProThrPheThrIleThrGlnIle	1480	QY	1821	AlaSerGlnIleAlaProProAlaGlyAlaThrGlyPheValValSerGlyLeuValGly	1840
DB	4721	GCAGTCACTCAAGTTGTAGACTTCAGTTTGGACCCGACCTTCATATAACACACAGACC	4780	DB	5801	GGGTCCCAATCGCGCACCCCGACGGGCCACTTGGCTTCTGTCTAGTGGCTGTGGGG	5860
QY	1481	ValProGlnAspAlaValSerArgSerGlnArgArgGlyValThrGlyArgGlyArgLeu	1500	QY	1841	AlaAlaValGlySerIleGlyLeuGlyLysValLeuValAspIleLeuAlaGlyTyrGly	1860
DB	4781	GTCCCTCAAGACCGCGTCTCAGCTAGCACCGCGCGGGGTGACGGGTAGGGGAAGACTG	4840	DB	5861	GCTGCGGTAGGAGCATAGGCTTGGTAAAGTGTGGTGGACATCTCTGCGAGGATGTG	5920
QY	1501	GlyIleTyrArgTyrValSerThrGlyGluArgAlaSerGlyMetPheAspSerValVal	1520	QY	1861	AlaGlyIleSerGlyAlaLeuValAlaPheLysIleMetSerGlyLysProSerMet	1880
DB	4841	GGCATTTATAGGTACGTGTCTGAGGAGGAGCTCGGGAATGTTTGACAGGTAGTG	4900	DB	5921	GGGGGATCTCGGGAGCTCTCGTCTGCAATTCAGATCATGTCTGGCGAGAGCCCTCATG	5980
QY	1521	LeuCysGluCysTyrAspAlaGlyAlaIleTyrTyrGluLeuThrProSerGluThrThr	1540	QY	1881	GluAspValValAsnLeuLeuProGlyIleLeuSerProGlyValAlaLeuValValGlyVal	1900
DB	4901	CTCAGTCAGTGTACGACTCAGGGCGCGCATGTGTACGAGCTCACACGCGGAGACACC	4960	DB	5981	GAGGAGCTCGTCAACCTGCTGCTGGAATCTCTCTCGGGTCTTTGGTGTGGGAGTCT	6040
QY	1541	ValArgLeuArgAlaTyrPheAsnThrProGlyLeuProValCysGlnAspHisLeuGlu	1560	QY	1901	IleCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrpMet	1920
DB	4961	GTGAGGCTCAGAGGCTATCTCAACACACACCTTGGCTTGGCTGTGGCCAAAGACCATCTGAG	5020	DB	6041	ATCTGCGCGGCCATCTACCGCCGACAGCTGGGACCGGGGAGGCGAGCTCCCAATGGATG	6100
QY	1561	PheTrpGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThr	1580	QY	1921	AsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaProThrHisTyrValThr	1940
DB	5021	TTTTGGAGGAGGATTTTACCGGCTTACACACATAGATGCCACCTTCTTTTCCCAACA	5080	DB	6101	AACAGGCTATTCCTCTCGCTTCCAGAGGAAACACAGTCTCGCCGCCACCCACTACGTGACG	6160
QY	1581	LysGlnSerGlyGluAsnAlaTyrLeuThrAlaTyrGlnAlaThrValCysAlaArg	1600	QY	1941	GluSerAspAlaSerGlnArgValThrGlnLeuLeuGlySerLeuThrIleThrSerLeu	1960
DB	5081	AAGCAAGCGGGGACAAATTCGCATACCTAACAGCCTTACCAGGCTACAGTGTGTGCCAGG	5140	DB	6161	GAGTCGATGCGTGGCGCGGTGACCCCACTACTCGGCTCCCTTACCATACTAGTCTG	6220
QY	1601	AlaLysAlaProProSerTrpAspValMetTrpLysCysLeuThrArgLeuLysPro	1620	QY	1961	LeuArgArgLeuHisAsnTrpIleThrGluAspCysProIleProCysGlyGlySerTrp	1980
DB	5141	GCCAAAGCCCTCCGTCGTCTGGGAGCTCATGTGGAAAGTGTGTGCTGACCTCAAGCCC	5200	DB	6221	CTTAGAAGACTCCACAAATTTGGATTACCGAGGGCTGCCCATCCCGTGACAGCGCTCGTGG	6280
QY	1621	ThrLeuValGlyProThrProLeuLeuTyrArgLeuGlySerValThrAsnGluValThr	1640	QY	1981	LeuArgAspValTrpAspTrpValCysThrIleLeuThrAspPheLysAsnTrpLeuThr	2000
DB	5201	ACGCTCGTGGCCCTTACCTCTCTGTACCGTGTGGCCCTGTCAACACAGAGGTCACT	5260	DB	6281	CTCCGCGATGTATGGGACTTGGGTCTGCACCATCTTAACAGACTTTTAAAACTGGCTGACC	6340
QY	1641	LeuThrHisProValThrLysTyrIleAlaThrCysMetGlnAlaLeuLeuValMet	1660	QY	2001	SerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTyrLys	2020
DB	5261	CTCACATCTCTGTGACGAAATACATCGCCACTTGCCATGCAAGCCGACCTTGAGGTCAATG	5320	DB	6341	TCTAAATTTGTCCAAAGATGCTGGCCCTCCCTTTGTCTCTTGCCAAAGGGGTACAAG	6400
QY	1661	ThrSerThrTrpValLeuAlaGlyGlyValLeuAlaAlaValAlaTyrCysLeuAla	1680	QY	2021	GlyValTrpAlaGlyThrGlyIleMetThrThrArgCysProCysGlyAlaAsnIleSer	2040
DB	5321	ACCAAGACGTGGGTCTTGGCTGGGGAGTCTTGGCGCCGCTGCTGCTGTATGCTTAGCG	5380	DB	6401	GGGTATGGCGCGGACTGTATCATGACACACAGGTGTCTTGTGCGGCCAATAATCTCT	6460
QY	1681	ThrGlyCysValCysIleIleGlyArgLeuHisIleAsnGlnArgAlaValAlaPro	1700	QY	2041	GlyAsnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTrp	2060
DB	5381	ACCGGTGTGTTTCATCATCCGCGCTTGACATCAACACAGAGGTGTCTGTGACCA	5440	DB	6461	GGCAATGTCCGCTGGGTCCATAGAAATCACAGGGCCCTAAGACCTGCATGAACACCTGG	6520
QY	1701	AspLysGluValLeuTyrGluAlaPheAspGluMetGluGluCysAlaSerArgAlaAla	1720	QY	2061	GlnGlyThrPheProIleAsnCysTyrThrGluGlyGlnCysValProLysProAlaPro	2080
DB	5441	GACAAAGAGGTCTCTATAGGCTTTTGATGAGATGGAGGAGTGTGCTCTAGAGCGGCC	5500	DB	6521	CAAGGAGACCTTTCCCATCAATTTGCTACACGGGGGGCGAGTGTGTGCGAAGCCCGGCCA	6580
QY	1721	LeuIleGluGlyGlnArgIleAlaGluMetLeuLysSerIleGlnGlyLeuLeu	1740	QY	2081	AsnPheLysValAlaIleTrpArgValAlaAlaSerGluTyrAlaGluValThrGlnHis	2100
DB	5501	CTCATTAAGAGGGGCGGATAGCCGAGATGTGAAGTCCAAATCCAAAGGTCTATTG	5560	DB	6581	AACTTCAAGACCCCATCTGGAGGTGGCGGCTCAGAGTACGCGAGGTGACGCAACAC	6640
QY	1741	GlnGlnAlaSerLysGlnAlaGlnAspIleGlnProThrValGlnAlaSerTrpProLys	1760	QY	2101	GlySerTyrHisTyrIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeu	2120
DB	5561	CAGCAGGCTTCAACAGCCACAGACATACAAACCCCGCTGTGAGGCTTCATGGCCCCAAG	5620	DB	6641	GGGTCACTCTCTATATAACAGGACTGACCAAGCAACCTGAAAGTTCCTTCCCAACTA	6700
QY	1761	ValGluGlnPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeuAla	1780	QY	2121	ProSerProGluPhePheSerTrpValAspGlyValGlnIleHisArgPheAlaProThr	2140
DB	5621	GTGAGGAGGATCTGGGGCAACACATGTGGAACTTCACTAGCGGCATTCATACCTTGCA	5680	DB	6701	CTTCTCCAGAGTCTTTTCTCTGGGTGGAGGAGTGCAGATCCACAGGTTCGCCCCCA	6760
QY				QY	2141	ProLysProPhePheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal	2160

REFERENCE/DOCKET NUMBER: 06/87-48009

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 659-2811

TELEFAX: (202) 659-1462

TELEX: WUI 64470

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 9589 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: single

TOPOLOGY: linear

US-07-925-695-1

Alignment Scores:

Pred. No.: 0 Length: 9589
Score: 15824.00 Matches: 2966
Percent Similarity: 98.78% Conservativity: 30
Best Local Similarity: 97.79% Mismatches: 37
Query Match: 98.16% Indels: 0
DB: 1 Gaps: 0

US-09-980-559-2 (1-3033) x US-07-925-695-1 (1-9589)

Qy 1 MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgProGln 20
Db 341 AUGAGCACAAUCCUAAACCCUAAAGAAAAACCAAAAGAAAAACCAACACCGGCGCCACAA 400
Qy 21 AspValLysPheProGlyGlyGlnValGlyGlyValLysValLysValLysValLysVal 40
Db 401 GACGUUAAAGUUUCCGGCGCGCGCGCAAGAUUGGCGGAGUAUUCUUGGCGCGCGAGG 460
Qy 41 GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly 60
Db 461 GCGCCAGGUG 520
Qy 61 ArgArgGlnProLysPheProLysPheProLysPheProLysPheProLysPheProLys 80
Db 521 AGGCGCGAGCCCAUCCUAAAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 580
Qy 81 TyrProTrpProLeuTyrGlyValGlnGlyValGlnGlyValGlnGlyValGlnGlyVal 100
Db 581 UACCCCGGCG 640
Qy 101 ArgGlySerArgProSerTrpGlyProAsnAspProArgHisArgSerArgAsnValGly 120
Db 641 CGAGGUUCCGUGUCCUUCUUGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 700
Qy 121 LysValLysAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrTrpValVal 140
Db 701 AAGGUCAUCGAUACCCUAAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 760
Qy 141 GlyAlaProLeuGlyGlyValAlaAlaArgAlaLeuAlaHisGlyValArgValLeuGlyAsp 160
Db 761 GGCAGCG 820
Qy 161 GlyValAsnPheAlaThrGlyValAsnLeuProGlyCysSerPheSerLysPheLeuLeuAla 180
Db 821 GGGGUUAAUUUUGCAACAGGGAACUUAACCGGUUGUCCUUCUUCUUCUUCUUCUUC 880
Qy 181 LeuLeuSerCysLysThrThrProValSerAlaAlaGluValLysAsnLysSerThrGly 200
Db 881 CUGCUGUCCUGCAUACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 940
Qy 201 TyrMetValThrAsnAspCysThrAsnAspSerLysThrTrpGlnLeuGlnAlaVal 220
Db 941 UACAGGUGAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1000
Qy 221 LeuHisValProGlyCysValProCysGluLysValGlyValAsnAlaSerGlnCysTrpIle 240
Db 1001 CUCCAGUCCCGGCG 1060
Qy 241 ProValSerProAsnValAlaValGlnArgProGlyAlaLeuThrGlnGlyLeuArgThr 260

Db 1061 CCGGUCUCCAGAAUGUGCGCGUGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1120
Qy 261 HisIleAspMetValValMetSerAlaThrLeuCysSerAlaLeuTyrValGlyAspLeu 280
Db 1121 CACAUUGACAUUGUUGAUGUCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1180
Qy 281 CysGlyGlyValMetLeuAlaAlaGlnMetPheIleValSerProGlnHisTrpPhe 300
Db 1181 UGCGGUGGUGUGAUGUCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1240
Qy 301 ValGlnAspCysAsnCysSerIleTyrProGlyThrIleThrGlyHisArgMetAlaTrp 320
Db 1241 GUGCAAGACUGCAUUGUCCCAUCCCGUGGUAUCCAUUGGCGCGCGCGCGCGCGCG 1300
Qy 321 AspMetMetMetAsnTrpSerProThrAlaThrMetIleLeuAlaTyrAlaMetArgVal 340
Db 1301 GACAUUGAUGAUGAUGUGUGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1360
Qy 341 ProGluValIleIleAspIleIleSerGlyAlaHisTrpGlyValMetPheGlyLeuAla 360
Db 1361 CCGGAGGUCAUCAUAGACAUCAUUGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1420
Qy 361 TyrPheSerMetGlnGlyAlaTrpAlaLysValValIleLeuLeuLeuLeuAlaGly 380
Db 1421 UACUUCUCUUAUGAGGAGCGUGGCGAAAGUGUGUUCUUCUUCUUCUUCUUCUUCG 1480
Qy 381 ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr 400
Db 1481 GUGAGCGCGCAAAACCAUACCGUGGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 1540
Qy 401 SerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTrp 420
Db 1541 GGCAUGUUCUCCUUGUGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1600
Qy 421 HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisTrpGlyPheIleAlaSer 440
Db 1601 CACAUCAAACCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1660
Qy 441 LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg 460
Db 1661 CUGUUCUACACCCACAGCUUCAAACUGUCAGGAGUCCCGGCGCGCGCGCGCGCG 1720
Qy 461 SerIleGluAlaPheArgValGlyTrpGlyAlaLeuGlnTyrGluAsnValThrAsn 480
Db 1721 AGUAUCGAGCGCUUCUGGUGGUGGUGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1780
Qy 481 ProGluAspMetArgProTyrCysTrpHisTyrProProArgGlnCysGlyValValSer 500
Db 1781 CCAGAGGAUAGAGACCGCUUUGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1840
Qy 501 AlaLysThrValCysGlyProValTyrCysPheThrProSerProValValValGlyThr 520
Db 1841 GCGAGCUCUGUGUGUGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1900
Qy 521 ThrAspArgLeuGlyAlaProThrThrThrTrpGlyGluAsnGluThrAspValPheLeu 540
Db 1901 ACCGAUAGACUUGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1960
Qy 541 LeuAsnSerThrArgProProLeuGlySerTrpPheGlyCysThrTrpMetAsnSer 560
Db 1961 UUGAAACAGCACUCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2020
Qy 561 GlyTyrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer 580
Db 2021 GGUUACACACAGACUUGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2080
Qy 581 ThrAspLeuLeuCysProThrAspCysPheArgLysHisPheProAspThrThrTyrLeuLys 600
Db 2081 AUGGACUUGUGUGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2140
Qy 601 CysGlySerGlyProTrpLeuThrProArgCysLeuLeuAspTyrProTyrArgLeuTrp 620
Db 2141 UGUGGUCUUGGCG 2200

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 11, 2004, 23:18:54 ; Search time 318 Seconds
(without alignments)
4209.796 Million cell updates/sec

Title: US-09-980-559-2
Perfect score: 16120
Sequence: 1 MSTNPKPQKRTNRRPQ.....LLFGLLLFVGGLFLLPAR 3033

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DSV=xlh
-Q=/cgn2_1/USFio.spool/US09980559/runat_09022004_141955_26421/app_query.fasta_1.3207
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09980559 @CGN 1 1 176 @runat_09022004_141955_26421 -NCPU=6 -ICPU=3
-NM MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEW TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15824	98.2	9589	1	US-07-925-695-1
2	15824	98.2	9589	1	US-07-925-695-2
3	13810	85.7	9511	1	US-07-925-695-7
4	13805	85.6	9511	1	US-07-925-695-6
5	11984.5	74.3	9030	1	US-08-324-977-13
6	11984.5	74.3	9030	2	US-08-384-616-13
7	11984.5	74.3	9030	2	US-08-904-686A-13
8	11984.5	74.3	9030	3	US-09-315-850-13
9	11984.5	74.3	9416	1	US-08-324-977-1
10	11984.5	74.3	9416	2	US-08-384-616-1
11	11984.5	74.3	9416	2	US-08-904-686A-1
12	11984.5	74.3	9416	3	US-09-315-850-1

ALIGNMENTS

RESULT 1
US-07-925-695-1
; Sequence 1, Application US/07925695
; Patent No. 5428145
; GENERAL INFORMATION:
; APPLICANT: OKAMURA, Hiroaki
; APPLICANT: OKAMURA, Tetsuo
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOME,
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, ANTIGEN, ANTIBODY AND
; TITLE OF INVENTION: DETECTION SYSTEMS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Beveridge, DeGrandi, Weilacher & Young
; STREET: 1850 M Street, N.W., Suite 800
; CITY: Washington
; STATE: D.C.
; COUNTRY: US
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/925,695
; FILING DATE: 19920807
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 287402/91
; FILING DATE: 09-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 360441/91
; FILING DATE: 05-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Weilacher, Robert G.
; REGISTRATION NUMBER: 20,531


```
Qy 2635 PheGlnTyrSerProAlaGlnArgValGluPheLeuLeuLysAlaTrpAlaGluLysLys 2654
Db 8175 TTCCAATACTCTCTCGACAGCGGTGCGAGTTCTCTGTGAATCCCTGGAAAGCAAGAAA 8234
Qy 2655 AspProMetGlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluArgAsp 2674
Db 8235 TGCCTATGGCTTCGCATATGACACCCGCTGTTTGACTCAACGGTCACTGAGATGAC 8294
Qy 2675 IleArgThrGluGluSerIleTyrArgAlaCysSerLeuProGluGluAlaHisThrAla 2694
Db 8295 ATCGGTGAGGAGTCAATCTACCAATGTGTGACTTGGCCCCCGAAGCCACAGGCC 8354
Qy 2695 IleHisSerLeuThrGluArgLeuTyrValGlyGlyProMetPheAsnSerLysGlyGln 2714
Db 8355 ATAAGTTCGCTCACAGAGCGCTTTACATCGGGGGCCCCCTGACTAATTTCTAAAGGGCAG 8414
Qy 2715 ThrCysGlyTyrArgArgCysArgAlaSerGlyValLeuThrThrSerMetGlyAsnThr 2734
Db 8415 AACTGCGGCTATCGCGGTGCGCGAGCGGTGTACTGACGACCACTGCGGTAAATACC 8474
Qy 2735 IleThrCysTyrValLysAlaLeuAlaCysLysAlaAlaGlyIleIleAlaProThr 2754
Db 8475 CTCACATGTACTTGAAGCGCGCTGCGGCTGTGAGCTCGAAGCTCCAGGACTGCACG 8534
Qy 2755 MetLeuValCysGlyAspLeuValIleSerGluSerGlnGlyThrGluGluAsp 2774
Db 8535 ATGCTGATGCGGAGACGACCTTGTGCTTATCTGTGAAAGCGCGGGGCCCAAGAGGAC 8594
Qy 2775 GluArgAsnLeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAsp 2794
Db 8595 GAGCGCAGCCTACGGGCTTTCACGGAGGCTATGACTAGATACTCTGCCCGCCCTGGGAC 8654
Qy 2795 ProProArgProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerVal 2814
Db 8655 CCGCCCAACCAAGATACGACTTGGAGTTGATAACATCATGCTCTCCCAATGTGTCAGT 8714
Qy 2815 AlaLeuGlyProGlnArgArgArgTyrTyrLeuThrArgAspProThrThrProIle 2834
Db 8715 GCGCAGATGCATCTGGCAAAAGGGTGTACTATCTCACCGGTGACCCACCCCCCTT 8774
Qy 2835 AlaArgAlaAlaTrpGluThrValArgHisSerProValAsnSerTrpLeuGlyAsnIle 2854
Db 8775 GCGCGGCTCGTGGGAGACAGCTAGACACACTCCAGTCAATTCCTGGCTAGGCAACATC 8834
Qy 2855 IleGlnTyrAlaProThrIleTrpAlaArgMetValLeuMetThrHisPhePheSerIle 2874
Db 8835 ATCATGTGCGGCCACCTTGTGGCAAGATGATCTCTGATGACTCATTTCTTCTCCATC 8894
Qy 2875 LeuMetAlaGlnAspThrLeuAspGlnAsnLeuAsnPheGluMetTyrGlyAlaValTyr 2894
Db 8895 CTTCTAGCTCAGGAACAACCTTGAAAGGCCCTAGATTGTGAGATCTACGGGGCTGTAC 8954
Qy 2895 SerValSerProLeuAspLeuProAlaIleIleGluArgLeuHisGlyLeuAspAlaPhe 2914
Db 8955 TCCATTGAGCCACTTGACCTTACCTCAGATCATTCACAGACTCCCATGGCTTAGCGCATTT 9014
Qy 2915 SerLeuHisThrTyrThrProHisGluLeuThrArgValAlaSerAlaLeuArgLysLeu 2934
Db 9015 TCACCTCATAGTTACTCTCAGGTGAGATCAATAGGGTGGCTTCATGCTCAGGAAACTT 9074
Qy 2935 GlyAlaProProLeuArgAlaTrpLysSerArgAlaArgAlaValArgAlaSerLeuIle 2954
Db 9075 GGGGTACCGCCCTTGGAGTCTGGAGACATCGGCGCCAGAGTGTCCGCGCTAGGCTACTG 9134
Qy 2955 SerArgGlyGlyArgAlaAlaValCysGlyArgTyrLeuPheAsnTrpAlaValLysThr 2974
Db 9135 TCCACAGGGGGAGGCTGCGACTTGTGGCAAGTACTCTTCAACTGGGCGAGTAAGGACC 9194
Qy 2975 LysLeuLysLeuThrProLeuProGluAlaArgLeuLeuAspLeuSerTrpPheThr 2994
Db 9195 AAGCTCAAACTCACTCAATCCCGCTGCGTCCAGTTGGATTTATCCAGCTGGTTTCGTT 9254
Qy 2995 ValGlyAlaGlyGlyAspIleTyrHisSerValSerArgAlaArgProArgLeuLeu 3014
```

```
Db 9255 GCTGGTTACAGCGGGGGAGACATATATCACAGCTGTCTCGTCCCGACCCCGCTGTTTC 9314
Qy 3015 LeuPheGlyLeuLeuLeuLeuPheValGlyValGlyLeuPheLeuLeuProAlaArg 3033
Db 9315 ATGTGGTGCCTACTCTCTACTTCTGTAGGGGTAGGCATCTATCTACTCTCCCAACCGA 9371
```

Search completed: February 12, 2004, 00:24:59
Job time : 3898 secs

6030 GTGTGCGCAGCATCTGGTTCGGCACGTGGGCCAGGCGGGGGCTGTGTCAGTGGATG 6089
1921 AsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaProThrHisTyrValThr 1940
6090 AACCGGCTGATAGCGTTCGCGGGGTAAACAGCTCTCCCAACGACACTATGTGCTT 6149
1941 GluSerAspAlaSerGlnArgValThrGlnLeuLeuGlySerLeuThrIleThrSerLeu 1960
6150 GAGAGCGAGCTGCACGAGTGTCACTCAGATCTCTCTAGTCTTACCATCACTCAGCTG 6209
1961 LeuArgArgLeuHisAsnTrpIleThrGluAlaPheProIleProCysGlyGlySerTrp 1980
6210 CTGAGAGGCTTCCACAGTGGATCAACAGGAGTGTCTCCACGCGCATCTCCGCTCGTG 6269
1981 LeuArgAspValTrpAspTrpValCysThrIleLeuThrAspPheLeuAsnTrpLeuThr 2000
6270 CTAAAGAGATGTTGGGATTTGGATATGACGGTGTGACTGATTTCAAGACCTGGCTCCAG 6329
2001 SerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTyrLys 2020
6330 TCCAAGCTCTGCGCGGATTCGCGGAGTCCCTTCTCTCATGTCAACGTGGGTACAAG 6389
2021 GlyValTrpAlaGlyThrGlyIleMetThrThrArgCysProCysGlyAlaLeuIleSer 2040
6390 GGAGTCTGGCGGGCGACGCGCATCATGCAACCAACCTGCCATGTGGAGCACAGATCAC 6449
2041 GlyAsnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTrp 2060
6450 GGACATGTGMAAACGGTTCATGAGATCTGGGGCTTAGGACCTGTAGTACACGTGG 6509
2061 GlnGlyThrPheProIleAsnCysTyrThrGluGlyGlnCysValProLysProAlaPro 2080
6510 CATGGAACATTCCTTAAACGGGTACACCAAGCGGCCCTGCACGCGCTCCCGCGGCCA 6569
2081 AsnPheLysValAlaIleTrpArgValAlaAlaSerGluTyrAlaGluValThrGlnHis 2100
6570 AATTATTCTAGGCGCTGTGGCGGGTGGTGTCTGAGAGTACGTGGAGGTTACGCGGGT 6629
2101 GlySerTyrHisTyrIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeu 2120
6630 GGGGATTTCCATGACGTGAGCGGCATGACCACTGACAACTGAAGTCCCGTGTACGGTT 6689
2121 ProSerProGluPhePheSerTrpValAspGlyValGlnIleHisArgPheAlaProThr 2140
6690 CCGGCCCCCAATCTTCACAGAGTGGATGGGTGGGTGGCACAGGTACGCTCCAGCG 6749
2141 ProLysProPhePheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal 2160
6750 TGCAAAACCCCTCTACGGGAGGAGTCAATCTCTGTCGGGCTCAATCAATACCTGGTT 6809
2161 GlySerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThr 2180
6810 GGGTTCACAGCTCCATGCGAGCGCGAACCAGGAGTAGCAGTGTCTCATCTTCCATGCTCACC 6869
2181 AspProSerHisIleThrAlaGluThrAlaAlaArgArgLeuAlaArgGlySerProPro 2200
6870 GACCCCTCCACATTACGGCGGAGAGCGCTAAGGTAGGTGGCCAGGGGATCTCCCCC 6929
2201 SerGluAlaSerSerAlaSerGlnLeuSerAlaProSerIleuArgAlaThrCysThr 2220
6930 TCCTTGGCCAGCTCATCAGCTAGCAGCTGTCTGCGCTTCTTGAAGGCAACATGCAC 6989
2221 ThrHisGlyLysAlaTyrAspValAspMetValAspAlaAsnLeuPhe----- 2236
6990 ACCGCTCATGCTCCCGGAGCGTCACTCATGAGGCCAACCTCTCTGTGGCGGAGGAG 7049
2237 MetGlyGlyAspValThrArgIleGluSerLysSerLysValValValLeuAspSerLeu 2256
7050 ATGGCGGGGAACATCACCCCGGTGGAGTCAAGAAATAAGTAGTAATTTTGGACTCTTTC 7109
2257 AspProMetValGluGluArgSerAspLeuGluProSerIleProSerGluTyrMetLeu 2276
7110 GAGCCGCTCCAAAGCGGAGGAGTACGAGGAGATATCCGTTCCGCGGAGATCTCTCGG 7169

2277 ProLysLysArgPheProProAlaLeuProAlaTrpAlaArgProAspTyrAsnProPro 2296
7170 AGGTCCAGGAATTCCTCGAGCGATGCCATATGGGCACGCCCGGATTACAACCTCCA 7229
2297 LeuValGluSerTrpLysArgProAspTyrGlnProAlaThrValAlaGlyCysAlaLeu 2316
7230 CTGTTAGAGTCTCTGGAAGGACCGGAGTACGCTCTCCAGTGGTACACGGGTGTCATTTG 7289
2317 ProProArgLysThrProThrProProArgArgArgArgThrValGlyLeuSer 2336
7290 CCGCTTGCACAGGCCCTCCGATACACTCCACGAGGAGGAGGACGGTGTCTCTGTCA 7349
2337 GluAspSerIleGlyAspAlaLeuGlnLeuAlaIleLysSerPheGlyGlnProPro 2356
7350 GAATCTACGCTCTCTTCTGCTTGGCGAGCTCGCCACAAAGACCTTCGGCAGCTCCGAA 7409
2357 ProSerGly---AspSerGlyLeuSerThrGlyAlaGlyAlaAlaAspSerGlySerGln 2375
7410 TCGTCGGCGCTCGACAGCGGCACGCAACGGCC-----TCTCTGACCCAG 7454
2376 ThrProProAspGluLeuAlaLeuSerGluThrThrGlySerIleSerSerMetProProLeu 2395
7455 CCTTCGACGACGCGGATCCGACGTTGAGTCTCTCTCCATGCCCCCTT 7514
2396 GluGlyGluLeuGlyAspProAspLeuGluProGluGlnValGluProGlnProProPro 2415
7515 GAGGGGAGCGGGGATCCCGATCTC----- 7541
2416 GlnGlyGlyValAlaAlaProGlySerAspSerGlySerTrpSerThrCysSerGluGlu 2435
7542 -----AGCGAGCGGTCTTGGCTTACCGTAAGCGAGGAG 7574
2436 ---AspAspSerValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuIleThr 2454
7575 GCTAGTAGGACGCTGTCTGCTGCTGATCTCTACATGACATGACAGGCGCTTGTATCAG 7634
2455 ProCysSerProGluGluLysLeuProIleAsnProLeuSerAsnSerLeuLeuArg 2474
7635 CCATGCGCTCGGAGGAAACCAAGCTGCCCATCAATGCACTGAGCAACTCTTGTCTCCGT 7694
2475 TyrHisAsnLysValTyrCysThrThrThrLysSerAlaSerLeuAlaLysLysVal 2494
7695 CACCACAACTTGTCTATGCTACAACTCTCGCAGCGCAAGCTTCGGCAGAGAGGTC 7754
2495 ThrPheAspArgMetGlnValLeuAspSerTyrTyrAspSerValLeuLysAspIleLys 2514
7755 ACCTTTGACAGACTGCGAGTCTTGACACCACTACCGGAGCGTGTCTCAAGGAGATGAAG 7814
2515 LeuAlaAlaSerLysValThrAlaArgLeuLeuThrMetGluGluAlaCysGlnLeuThr 2534
7815 GCGAGGCGCTCCACAGTTAAGGCTTAACTTCTATCGTGGAGAGCCCTGTAAGCTGACG 7874
2535 ProProHisSerAlaArgSerLysTyrGlyPheGlyAlaLysGluValArgSerLeuSer 2554
7875 CCCCCACATTCGCCAGATCTAAATTTGGCTATGGGGCAAGAGCGTCCGGAACTATCC 7934
2555 GlyArgAlaValAsnHisIleLysSerValTrpLysAspLeuLeuGluAspSerGluThr 2574
7935 AGCAAGCGGCTTAACACATCCGCTCCGTGTGAAGGACTTGTCTGGAAGACACTGAGACA 7994
2575 ProIleProThrThrIleMetAlaLysAsnGluValPheCysValAspProThrLysGly 2594
7995 CCAATTTGACACCACCATCATGGCAAAATAGGTTTTCTCGCTCCCAACAGAGAGGGG 8054
2595 GlyLysAlaAlaArgLeuIleValTyrProAspLeuGlyValArgValCysGluLys 2614
8055 GCGCGAAGCGACTCGCTCTATCGTATTCAGATTTGGGGTTCGTGTGTCGAGAA 8114
2615 MetAlaLeuTyrAspIleThrGlnLysLeuProGlnAlaValMetGlyAlaSerTyrGly 2634
8115 ATGGCCCTTTACGATGTGTCTCCACCTCTCCAGCGCTGATGGGCTCTTCATACGGA 8174

[illegible]

1662 CTGTTCTAGCTGCACAAAGTTCAACTCATCTGGATGCCAGAGCGCATGGCCAGCTGCAGC 1721
 461 SerIleGluAlaPheArgValGlyTTPGlyAlaLeuGluTyrGluAspAsnValThrAsn 480
 1722 CCATCGAGCGGTTCGCTCAGGGGTGGGGCCCATCACTTAC-----AATGAGTCACAC 1775
 481 ProGluAspMetArgProTyrCysTrpHisTyrProProArgGlnCysGlyValValSer 500
 1776 AGCTCGGACGAGGCGCTATTGTTGGCACTACGCACCCCGCGCGGTATGTCATGCC 1835
 501 AlaLysThrValCysGlyProValTyrCysPheThrProSerProValValValGlyThr 520
 1836 GCGGCGAGGTGTGTGTCAGTGTACTGCTTCCACCCCAAGCGCTGTGCTGGTGGGAGC 1895
 521 ThrAspArgLeuGlyAlaProThrTyrThrTyrGlyGluAsnGluThrAspValPheLeu 540
 1896 ACCGACCGGTTCGCGCTCCCTACGTACGTTGGGGGAGNATGAGACGACGTGCTGCTT 1955
 541 LeuAsnSerThrArgProProLeuGlySerTrpPheGlyCysThrTrpMetAsnSerSer 560
 1956 CTTTAACAACACGCGCGCGCAAGCAACTGTTGGCTGTATCATGGATGAATAGCACT 2015
 561 GlyTyrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer 580
 2016 GGGTTTCAACCAAGACGTGCGGGGGCGCGCGCTGTAAACATCGGGGGGATCGGCAATAAAACC 2075
 581 ThrAspLeuLeuCysProThrAspCysPheArgGlyHisProAspThrThrTyrLeuLys 600
 2076 -----TTGACCTTGCCTCCACGAGTGTCTCCGAAAGCACCCCGAGGCCATTACCAAG 2129
 601 CysGlySerGlyProTrpLeuThrProArgCysLeuLeuAspTyrProTyrArgLeuTrp 620
 2130 TGTGGTTCGGGGCTTGGTTGACACCCAGATGCTTGGTCCACTACCCATACAGCTTGG 2189
 621 HisTyrProCysThrValAsnTyrThrIlePheLysIleArgMetTyrValGlyVal 640
 2190 CACTACCCCTGCACCTGTCAACTTACCATCTTCAAGGTAGGATGTACGTGGGGGAGTG 2249
 641 GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp 660
 2250 GAGCACAGGCTCGAAGCGCATGCATCAATTGACTCGAGGAGGCTGTGAACCTTGGAGGAC 2309
 661 ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGluTrpAlaIleLeuPro 680
 2310 AGGACAGATCAGAGCTTAGCCCGCTGCTGCTTACACAGGAGTGGCAGGTATTGCC 2369
 681 CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisLeuHisGlnAsnIle 700
 2370 TGTTCCTTCAACACCCCTACCGGCTCTGTCCACTGGTTTGTATCCATCTCCATCAGAAGTC 2429
 701 ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrLysTyrIleValArgTyr 720
 2430 GTGACGTCAATACCTGTACGGTATAGGCTGGCGGTTGTCTCTTTCGCAATCAATAGG 2489
 721 GluTrpValIleLeuLeuPheLeuLeuAlaAspAlaArgValCysAlaCysLeuTrp 740
 2490 GAGTATGCTCTGTGCTCTTCTTCTTGGGGACGCGCGCTGTGTGCTGTGCTGG 2549
 741 MetLeuIleLeuLeuGlyGlnAlaGluAlaLeuGluLysLeuValIleLeuHisAla 760
 2550 ATGATGCTGTGATAGCTCAAGCTGAGCGCGCTAGAGAACCTGTGTGTCTCCACGCG 2609
 761 AlaSerAlaAlaSerCysAsnGlyPheLeuTyrPheValIlePhePheValAlaAlaTyr 780
 2610 GCATCCGTGGCGGGGCGCATGGCATTTCTCTCTCTCTGCTGTCTTCTGTGCTGCTGG 2669
 781 TyrIleLysGlyArgValValProLeuAlaThrTyrSerLeuThrGlyLeuTrpSerPhe 800
 2670 TACATCAAGGCGAGCGTGGTCCCTTGGGGCGGCATATGCCCTCTACGGGCTATGGCCGCTA 2729
 801 SerLeuLeuLeuLeuAlaLeuProGlnGlnAlaTyrAlaTyrAspAlaSerValHisGly 820
 2730 CTCCTGCTCTGCTGGGCTTACCACCGAGCATACGCCATGGACCGGAGATGGCAGCA 2789

821 GlnIleGlyAlaAlaLeuLeuValMetIleThrLeuPheThrLeuThrProGlyTyrLys 840
 2790 TCGTCGGAGGCGCGTTCGTAGTCTGTACTCTTGGACCTTGTCCACCGCACTATAAG 2849
 841 ThrLeuLeuSerArgPheLeuTrpTrpLeuCysTyrLeuLeuThrLeuGlyGluAlaMet 860
 2850 CTGTCTCTCGCTAGCTCATATGTTGGTTCACAAATATTTATCACCAGGCGCGGAGCACAC 2909
 861 ValGlnGluTyrAlaProProMetGlnValArgGlyGlyArgAspGlyIleIleTrpAla 880
 2910 TTGCAAGTGTGATCCCCCTCAAGCTTCGGGGGGCGCGCATGTCGCTCATCTCTCTC 2969
 881 ValAlaIlePheTyrProGlyValValPheAspIleThrLysTrpLeuLeuAlaValLeu 900
 2970 ACGTCGGCGATCCACCCAGAGCTAATCTTACCATCACCATAATCTTGTCTGCCATATCTC 3029
 901 GlyProAlaTyrLeuLeuLysGlyAlaLeuThrArgValProTyrPheValArgAlaHis 920
 3030 GGTCCACTCATGTCCTCCAGCTGTATTAACCAAGTGCCTACTTCTGTGGCGGCACAC 3089
 921 AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyGlyArgTyrValGlnMet 940
 3090 GGGCTCATTCGTGCATGATGCTGTGCGAAGGTTCCTGGGGGTTCATTATGTCTCCAAATG 3149
 941 AlaLeuLeuAlaLeuGlyArgTyrThrGlyThrTyrIleTyrAspHisLeuThrProMet 960
 3150 GCTCTCATGAAGTTGGCCGCACTGACAGGTACGTAGCTTATGACCATCTCACCACCTG 3209
 961 SerAspTrpAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIleIlePhe 980
 3210 CGGAGCTGGCGCCACGCGGCTTACGAGACCTTGGGTGGGAGTTGAGCGCTGCTCTTC 3269
 981 SerProMetGluLysLysValIleValTyrGlyAlaGluThrAlaAlaCysGlyAspIle 1000
 3270 TCTGATATGGACCAAGGTTATCCTGGGGGGGAGACACCGCGCGTGTGGGACATC 3329
 1001 LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp 1020
 3330 ATCTTGGGCTGCTGCTCTCCCGCGAGGGGAGGAGATACATCTGGGACCGCGCAGAC 3389
 1021 GlyTyrThrSerLysGlyTrpSerLeuLeuAlaProIleThrAlaTyrAlaGlnGlnThr 1040
 3390 AGCCTTGAAGGCGGCGGTGGGAGCTCTCTCGGCGCTATTACGCGCTTACTCCCAACAGC 3449
 1041 ArgGlyLeuLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGluGlnAla 1060
 3450 CGAGGCTTACTTGGCTGTGCATCATCTAGCTCTACAGCGCGGAGACAGGACCGAGTGCAG 3509
 1061 GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly 1080
 3510 GGGGAGGTCCAAAGTGTCTCCACCGCAACACAATCTTCTTCTGCGACCTTGTCAATGGC 3569
 1081 ValLeuTrpThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro 1100
 3570 GTGTGTGGACTGTCTATCATGTGCGGCTCAAAGACCTTTCGCGCGCCCAAGGGGCCCA 3629
 1101 ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTyrProSerProGly 1120
 3630 ATCACCCTCAATGTACCAATGTGGACCAAGGAGCTGTGCGCTGGCAAGCGCCCCCGGG 3689
 1121 ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArgAsn 1140
 3690 GCGGCTTCTTGACACCATGTCACCTGCGGAGCTCGGACCTTTACTTGTTCAGGAGCAT 3749
 1141 AlaAspValIleProAlaArgArgGlyAspLysArgGlyAlaLeuLeuSerProArg 1160
 3750 GCGGATGTCTTCCGCGCGCGCGCGCGCGAGCAGCAGGAGGAGGAGCTTCTCTCCCGAG 3809
 1161 ProLeuSerThrLeuLysGlySerSerGlyGlyProValIleCysProArgGlyHisAla 1180
 3810 CCGCTCTCTTACTTGAAGGCTCTTCTGGCGGCTCTCTGCTCTGCTCTGCGGCGAGCT 3869

XX 01-AUG-2002. DB
 XX 16-JAN-2002; 2002WO-EP00526. QY
 XX 23-JAN-2001; 2001US-263479P. DB
 XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI. QY
 PA De Francesco R, Migliaccio G, Paonessa G; DB
 XX WPI; 2002-599793/64. QY
 XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV DB
 PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal QY
 PT ribosome entry site (IRES) region, useful in studying HCV replication DB
 PT and expression - QY
 XX Claim 9; Page -: 69pp; English. DB
 XX The invention relates to nucleic acid molecules comprising altered HCV QY
 CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV) DB
 CC internal ribosome entry site (IRES) region coding for one or more NS3, QY
 CC NS5A, or EMCV IRES mutations, respectively. The location of the DB
 CC mutations are detailed in the specification. Also included are QY
 CC (1) an expression vector comprising a nucleotide sequence coding for DB
 CC the altered nucleic acids, which is transcriptionally coupled to an QY
 CC exogenous promoter; (2) a recombinant cell human hepatoma cell comprising DB
 CC the altered nucleic acids; (3) a recombinant cell produced by introducing QY
 CC into a human hepatoma cell the altered nucleic acids; (4) producing an DB
 CC HCV (hepatitis C virus) replicon enhanced cell or which containing a QY
 CC functional HCV replicon; (5) an HCV replicon enhanced cells made in the DB
 CC method; and (6) measuring the ability of a compound to affect HCV QY
 CC activity. The HCV replicons and HCV replicon enhanced cells are useful in DB
 CC studying HCV replication and expression, and HCV and host cell QY
 CC interactions, producing HCV RNA and proteins, and providing a system DB
 CC for measuring the ability of a compound to modulate one or more HCV QY
 CC activities e.g. to discover drugs which may treat HCV mediated DB
 CC diseases such as liver failure, cirrhosis and hepatocellular carcinoma. QY
 CC The present sequence is an HCV replicon Con 1 mutant of the invention. DB
 CC Note: The present sequence is not shown in the specification but QY
 CC was created by the indexer using the HCV sequence appearing as DB
 CC ABK91411 and the information in Claim 9. QY
 XX Sequence 9605 BP; 1909 A; 2883 C; 2734 G; 2079 T; 0 other; SQ

Alignment Scores:
 Pred. No.: 0 Length: 9605
 Score: 12028.50 Matches: 2183
 Percent Similarity: 83.38% Conservative: 351
 Best Local Similarity: 71.83% Mismatches: 470
 Query Match: 74.62% Indels: 35
 DB: 24 Gaps: 7

US-09-980-559-2 (1-3033) x ABK91425 (1-9605)

QY 1 MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgProGln 20
 DB 342 ATGAGCAGCAATCTTAACCTCAAGAAAAACCAACGTAACCAACCGCCGCCACAG 401
 QY 21 AspValLysPheProGlyGlyGlnIleValGlyValLysValLysLeuProArgArg 40
 DB 402 GAGGTCAAGTTCCTGGGGCGGTGTCAGATCGTCGGTGGAGTTACCTGTTCGGCGCAGG 461
 QY 41 GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly 60
 DB 462 GCGCCCGAGGTGGGTGTGGCGCGCACTAGGAAGACTTCCGAGCGGTGCGCAACTCGTGA 521
 QY 61 ArgArgGlnProLysAspArgArgSerThrGlyLysSerThrProGlyLysProGly 80
 DB 522 AGCGCAACCTATCCCAAGGTCGCCAGCGCGGTAGGCGCTGGGCTCAGCCCGGG 581
 QY 81 TyrProTrpProLeuTyrGlyAsnGluGlyLeuGlyTyrAlaGlyTyrProLeuLeuSerPro 100

582 TACCCCTGGCCCTCTCTATGGCAATGAGGCTTGGGGTGGCGAGGATGGCTCTGTGTACCC 641
 QY 101 ArgGlySerArgProSerTrpGlyProhenAspProArgHisArgSerArgenValGly 120
 DB 642 CGTGGCTCTCGGCGCTAGTTGGGGCCCCACGGACCCCGCGGTAGGTTCGGCAATTTGGGT 701
 QY 121 LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProValVal 140
 DB 702 AAGGTCAATGATACCTCAGTGGGCTTCGCGGATCTCATGGGTACATTCGGCTCGTC 761
 QY 141 GlyAlaProLeuGlyGlyValAlaArgAlaLeuAlaHisGlyValArgValLeuGluApp 160
 DB 762 GCGCGCCCTTACGGGGCGCTGCCAGGGCCCTGGCGCATGGCTCCGGGTTCGGAGGAC 821
 QY 161 GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheIleAla 180
 DB 822 GCGGTGAATGATGCAACAGGGAATTCGCCGGTTCCTCTTCTATCTTCTTTGGCT 881
 QY 181 LeuLeuSerCysIleThrThrProValSerAlaAlaGluValLysAsnIleSerThrGly 200
 DB 882 TTGCTGTCTCTGTTGACCATCCAGCTTCGGTTCATGAGTGGCAACGATTCGGAGTG 941
 QY 201 TyrMetValThrAsnAspCysThrAsnAspSerIleThrTrpGlnLeuGlnAlaVal 220
 DB 942 TACCATGTCAAGCAACGACTGCTCCAAAGCAAGCATTTGTATGAGCGCAGGCATGATC 1001
 QY 221 LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTrpIle 240
 DB 1002 ATGCATACCCCGGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1061
 QY 241 ProValSerProAsnValAlaValGlnArgProGlyAlaLeuThrGlnGlyLeuArgThr 260
 DB 1062 GCGCTCATCTCCCGCTCGCGCCAGGACGCTAGGTCCTCCCTACGACGATACGACGC 1121
 QY 261 HisIleAspMetValValMetSerAlaThrLeuCysSerAlaLeuTyrValGlyAspLeu 280
 DB 1122 CATGTGCAATTTCTGTTGGGGCGGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1181
 QY 281 CysGlyGlyValMetLeuAlaGlnMetPheIleValSerProGlnHisIleStrpPhe 300
 DB 1182 TCGCGATCTGTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1241
 QY 301 ValGlnAspCysAsnCysSerIleTyrProGlyThrIleThrGlyHisArgMetAlaTyr 320
 DB 1242 GTACAGGACTGCAATTTGCTCAATATATCCCGCCACGTCAGGTCACGTCATGCTGCTG 1301
 QY 321 AspMetMetMetAsnTrpSerProThrAlaThrMetIleLeuAlaTyrAlaMetArgVal 340
 DB 1302 GATATGATGATGAATGCTGCTACAGCAGCCCTAGTGGTATCGCAGTTACTTCGGGATC 1361
 QY 341 ProGluValIleAspIleSerGlyAlaHisTrpGlyValMetPheGlyLeuAla 360
 DB 1362 CCACAAGCTGTGCGTATGTTGGTGGGGGCCCATTTGGGGAGTCTCTAGCGGGCTTGGC 1421
 QY 361 TyrPheSerMetGlnGlyAlaTrpAlaValValIleLeuLeuAlaGly 380
 DB 1422 TACTATTTCATGTCGGGAACCTGAGGCTTGAATGATGCTACTCTTTGCGGCG 1481
 QY 381 ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr 400
 DB 1482 GTTGAGGGGGAACCTATGTGACAGGGGGGAGTGGCCAAACACCCCTCGGGATTACG 1541
 QY 401 SerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTrp 420
 DB 1542 TCCCTCTTTTCCACCGGGGTTCATCCAGAAAAATCCAGTCTGTAAACACCAACGCGAGTGG 1601
 QY 421 HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer 440
 DB 1602 CACATCAACAGACTGCTGCTGAATGACTCCCTCAACACTGGGTTCCTTGTCTGG 1661
 QY 441 LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg 460

8655	CGCCCCAAACAGATACAGCACTTGGAGTTGATACATCATCTCTCTCAATGTGTGAGTC	8714	CGCCCCAAACAGATACAGCACTTGGAGTTGATACATCATCTCTCTCAATGTGTGAGTC
2815	AlaLeuGlyProGlnGlyArgArgTyrTyrLeuThrArgAspProThrThrProIle	2834	AlaLeuGlyProGlnGlyArgArgTyrTyrLeuThrArgAspProThrThrProIle
8715	CGCCAGCATGTCATCTGGCAAAAGGGTGTACTATCTCACCGGTGACCCACACCCGCCCTT	8774	CGCCAGCATGTCATCTGGCAAAAGGGTGTACTATCTCACCGGTGACCCACACCCGCCCTT
2835	AlaArgAlaAlaTyrGluThrValArgHisSerProValAsnSerTyrLeuGlyAsnIle	2854	AlaArgAlaAlaTyrGluThrValArgHisSerProValAsnSerTyrLeuGlyAsnIle
8775	GCAGCGGCTCGTGGGAGACAGCTAGACACACTCCAGTCAATTCCTGGCTAGGCAACATC	8834	GCAGCGGCTCGTGGGAGACAGCTAGACACACTCCAGTCAATTCCTGGCTAGGCAACATC
2855	IleGlnTyrAlaProThrIleTyrAlaArgMetValLeuMetThrHisPhePheSerIle	2874	IleGlnTyrAlaProThrIleTyrAlaArgMetValLeuMetThrHisPhePheSerIle
8835	ATCATGTATGGCCCACTTGTGGCGAAGGATGATCTCTGATGACTCATTTCTTCTCCATC	8894	ATCATGTATGGCCCACTTGTGGCGAAGGATGATCTCTGATGACTCATTTCTTCTCCATC
2875	LeuMetAlaGlnAspThrLeuAspGlnAsnLeuAsnPheGluMetTyrGlyAlaValTyr	2894	LeuMetAlaGlnAspThrLeuAspGlnAsnLeuAsnPheGluMetTyrGlyAlaValTyr
8895	CTTCTAGCTCAGGAACAACCTTGAAAAAGCCCTAGATTGTCAATCTACGGGGCTGTGTAC	8954	CTTCTAGCTCAGGAACAACCTTGAAAAAGCCCTAGATTGTCAATCTACGGGGCTGTGTAC
2895	SerValSerProLeuAspLeuProAlaIleIleGluArgLeuHisGlyLeuAspAlaPhe	2914	SerValSerProLeuAspLeuProAlaIleIleGluArgLeuHisGlyLeuAspAlaPhe
8955	TCCATTGAGCCACTTGACCTACCTCAGATCATTTCAACGACTCCATGGCCCTTAGCGCATTT	9014	TCCATTGAGCCACTTGACCTACCTCAGATCATTTCAACGACTCCATGGCCCTTAGCGCATTT
2915	SerLeuHisThrTyrThrProHisGluLeuThrArgValAlaSerAlaLeuArgLysLeu	2934	SerLeuHisThrTyrThrProHisGluLeuThrArgValAlaSerAlaLeuArgLysLeu
9015	TCATCTCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGCTTCATGCCTCAGGAACTT	9074	TCATCTCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGCTTCATGCCTCAGGAACTT
2935	GlyAlaProProLeuArgAlaTyrLysSerArgAlaArgAlaValArgAlaSerLeuIle	2954	GlyAlaProProLeuArgAlaTyrLysSerArgAlaArgAlaValArgAlaSerLeuIle
9075	GGGGTACCGCCCTTGGCGAGTCTGGAGACATCGGGCCAGAAGTGTCCGGCTAGGCTACTG	9134	GGGGTACCGCCCTTGGCGAGTCTGGAGACATCGGGCCAGAAGTGTCCGGCTAGGCTACTG
2955	SerArgGlyGlyArgAlaAlaValCysGlyArgTyrLeuPheAsnTyrAlaValLysThr	2974	SerArgGlyGlyArgAlaAlaValCysGlyArgTyrLeuPheAsnTyrAlaValLysThr
9135	TCCACAGGGGGAGGGGCTGCCACTTGTGGCAAGTACCTCTTCACTGGGCAGTAGGACC	9194	TCCACAGGGGGAGGGGCTGCCACTTGTGGCAAGTACCTCTTCACTGGGCAGTAGGACC
2975	LysLeuLysLeuThrProLeuProGluAlaArgLeuLeuAspLeuSerSerTyrPheThr	2994	LysLeuLysLeuThrProLeuProGluAlaArgLeuLeuAspLeuSerSerTyrPheThr
9195	AAGTCAAACTCATCTCCAATCCCGCTGGTCCAGTTGGATTTATCCAGCTGTTCTGTT	9254	AAGTCAAACTCATCTCCAATCCCGCTGGTCCAGTTGGATTTATCCAGCTGTTCTGTT
2995	ValGlyAlaGlyGlyGlyAspIleTyrHisSerValSerArgAlaArgProArgLeuLeu	3014	ValGlyAlaGlyGlyGlyAspIleTyrHisSerValSerArgAlaArgProArgLeuLeu
9255	GCTGTTACACGGGGGAGACATATATACAGACCTGTCTCGTGTCCGACCCCGCTGGTTC	9314	GCTGTTACACGGGGGAGACATATATACAGACCTGTCTCGTGTCCGACCCCGCTGGTTC
3015	LeuPheGlyLeuLeuLeuLeuPheValGlyValGlyLeuPheLeuLeuProAlaArg	3033	LeuPheGlyLeuLeuLeuLeuPheValGlyValGlyLeuPheLeuLeuProAlaArg
9315	ATGTGGTGCCTACTCTCTTCTGTAGGGGTAGGCATCTATCTACTTCCCCAACCCGA	9371	ATGTGGTGCCTACTCTCTTCTGTAGGGGTAGGCATCTATCTACTTCCCCAACCCGA

RESULT 15

ABK91425 standard; DNA; 9605 BP.

ABK91425;

15-NOV-2002 (first entry)

Hepatitis C virus Con 1 isolate DNA mutant 2.

HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis; hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV; internal ribosome entry site; IRES; NS5A; HCV replication; mutant.

Hepatitis C virus.

Synthetic.

Key	Location/Qualifiers
CDS	342..9374
	/*tag= a
	/product= "HCV polypeptide"
	/note= "The polypeptide consists of the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
mutation	replace (3946,A)
	/*tag= b

WO200259321-A2.

WO200259321-A2.

1721 LeuIleuGluGluGlnArgIleAlaGluMetLeuLysSerLysIleGlnGlyLeuLeu 1740
 5490 TACATCGAACAGGGAATGAGCTCGCGAACAATCAAAACAGAGGCAATCGGGTGTCTG 5549
 1741 GlnGlnAlaSerLysGlnAlaGlnAspIleGlnProThrValGlnAlaSerTrpProLys 1760
 5550 CAAACAGCCACCAAGCAGGAGGAGTGTCTCCGCTGGTGAATCCAAAGTGGCGGACC 5609
 1761 ValGluGlnPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTrpLeuAla 1780
 5610 CTCGAAGCCTTCGCGGAGAGCATATGTGGAATTTTCATAGCGGATACAAATATTAGCA 5669
 1781 GlyLeuSerThrLeuProGlyAsnProAlaValAlaSerMetMetAlaPheSerAlaAla 1800
 5670 GGCTGTTCACATCTGCTGCAACCCCGCATAGCATCACTGATGGCATTCACAGCTCT 5729
 1801 LeuThrSerProLeuSerThrThrThrIleLeuLeuAsnIleLeuGlyIleTrpLeu 1820
 5730 ATCACCAGCCGCTCACCAACATACCTCTCTTTAAACATCTCGGGGATGGGTG 5789
 1821 AlaSerGlnIleAlaProAlaGlyAlaThrGlyPheValSerGlyLeuValGly 1840
 5790 GCGCCCACTCTCTCCAGCGTCTCTGCTTCTGTAGCGCGGCATCGCTGGA 5849
 1841 AlaAlaValGlySerIleGlyLeuGlyLysValLeuValAspIleLeuAlaGlyIleGly 1860
 5850 GCGGCTGTGGCAGCATAGSCCTTGGGAGGTCTTGTGGATATTTGGCAGGTATGGA 5909
 1861 AlaGlyIleSerGlyAlaLeuValAlaPheLysIleMetSerGlyGluLysProSerMet 1880
 5910 GCAGGGGTGGCAGCGCGCTCTGTGGCTTTAAGGTCTAGCGCGGAGATGCCCTCCACC 5969
 1881 GluAspValValAsnLeuLeuProGlyIleLeuSerProGlyAlaLeuValGlyVal 1900
 5970 GAGGACCTGTGTTAACTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 6029
 1901 IleCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrpMet 1920
 6030 GTGTGGCGCAGCATCTGCTGGCAGCTGGGCGCCAGGGAGGGGGCTGTGCAGTGGATG 6089
 1921 AsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaProThrHisThrValThr 1940
 6090 AACCGGCTGATAGCTTCTGCTCGGGGTAAACACGCTCTCCGCCACGCACTATGTGCT 6149
 1941 GluSerAspAlaSerGlnArgValThrGlnLeuLeuGlySerLeuThrIleThrSerLeu 1960
 6150 GAGAGCGAGCTGCAGCAGCTGTCACTCAGATCTCTCTCTCTCTCTCTCTCTCTCTCT 6209
 1961 LeuArgArgLeuHisAsnTrpIleThrGluAspCysProIleProCysGlyGlySerTrp 1980
 6210 CTGAAGAGGCTTCACAGTGGATCAAGAGGAGTCTCCACGCCATGCTCCGGCTCGTGG 6269
 1981 LeuArgAspValTrpAspTrpValCysThrIleLeuThrAspPheLysAsnTrpLeuThr 2000
 6270 CTAAGAGATGTTGGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 6329
 2001 SerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTrpLys 2020
 6330 TCAACGCTCTGCGCGGATGGCGGAGTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 6389
 2021 GlyValTrpAlaGlyThrGlyIleMetThrThrArgCysProCysGlyAlaAsnIleSer 2040
 6390 GGAGTCTGGCGGCGCAGCGCATCATGCAACACCTGCCCATGTGGAGCAGCATCACC 6449
 2041 GlyAsnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTrp 2060
 6450 GGACATGTGAATAACGGTTCATGAGATCGTGGGGCTTAGGACCTGTAGTAACAGTGG 6509
 2061 GlnGlyThrPheProIleAsnCysThrThrGluGlyGlnCysValProLysProAlaPro 2080
 6510 CATGGAACATTCCTCCATTAACCGGTACACCAAGCGGGCCCTGCACGCGCTCTCCCGGCGCA 6569

2081 AsnPheLysValAlaIleTrpArgValAlaAlaSerGluTrpAlaGluValThrGlnHis 2100
 6570 AATTATTCTAGGCGCTGTGGGGTGGTGTCTGAGAGTACGTGAGGTTACGCGGTG 6629
 2101 GlySerTyHisTrpIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeu 2120
 6630 GGGGATTTCCATCTACGTGACGGGCATGACCACTGACCAAGTAAAGTCCCGTGTCAAGTT 6689
 2121 ProSerProGluPhePheSerTrpValAspGlyValGlnIleHisArgPheAlaProThr 2140
 6690 CCGGCCCCGAATTTCTACAGAAGTGGATGGGTGGTGTGACAGGTACGCTCCAGCG 6749
 2141 ProLysProPhePheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal 2160
 6750 TGCMAACCCCTCTACGGGAGGAGGTCAATTCCTGCTGGGTCAATCAATACCTGTT 6809
 2161 GlySerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThr 2180
 6810 GGGTCACAGTCCCATGCGAGCCGAAACCGAGCGTAGCAGTGTCTCACTTCCATGCTCACT 6869
 2181 AspProSerHisIleThrAlaGluThrAlaAlaArgAtcLeuAlaArgGlySerProPro 2200
 6870 GACCCCTCCCACTACGGCGGAGCGCTAAGCGTAGGTGGCCAGGGGATCTCCCCC 6929
 2201 SerGluAlaSerSerAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThr 2220
 6930 TCCTTGGCCAGCTCATCAGTAGCCAGCTGTCTGGCCTTCTTGAAGGCAACATGCACT 6989
 2221 ThrHisGlyLysAlaTrpAspValAspMetValAspAlaAsnLeuPhe 2236
 6990 ACCCGCTCATGACTCCCGGAGCGTCACTCATCGAGGCCAAGCTCTGTGGCGGAGGAG 7049
 2237 MetGlyGlyAspValThrArgIleGluSerGlySerLysValValValLeuAspSerLeu 2256
 7050 ATGGCGGGAACATCACCCGCTGGAGTCAGAAAATAAGTAGTAGTATTTTGGACTCTTC 7109
 2257 AspProMetValGluGluArgSerAspLeuGluProSerIleProSerGluTrpMetLeu 2276
 7110 GAGCGCTCCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 7169
 2277 ProLysLysArgPheProProAlaLeuProAlaLeuProAlaLeuProAlaLeuProPro 2296
 7170 AGGTCCAGGAAATTCCTCGAGCGATGCCATATGGGACGCGCGGATTAACAACCTCCA 7229
 2297 LeuValGluSerTrpLysArgProAspTrpGlnProAlaThrValAlaGlyCysAlaLeu 2316
 7230 CTGTTAGAGTCTCGAAGGACCCGAGTACGCTCTCCAGTGTGTACACGGGTGTCCATTG 7289
 2317 ProProProArgLysThrProThrProProProArgArgArgArgThrValGlyLeuSer 2336
 7290 CCGCTCCCAAGGCCCTCCGATACCACTCCACGAGGAGGAGGAGGAGGAGGAGGAGGAG 7349
 2337 GluAspSerIleGlyAspAlaLeuGlnLeuAlaIleLysSerPheGlyGlnProPro 2356
 7350 GAATCTACCGTCTCTCTGCTTGGCGAGTCCGACAAAGACCTTCGGCAGCTCCGAA 7409
 2357 ProSerGly---AspSerGlyLeuSerThrGlyAlaGlyAlaAlaAspSerGlySerGln 2375
 7410 TCGTCGGCGCTGCAGCGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACGCGGAC 7454
 2376 ThrProProAspGluLeuAlaLeuSerGluThrGlySerIleSerSerMetProLeu 2395
 7455 CCCTCCGAGCGGCGGAGTCCGAGCTGGAGTGTGAGTCTCTCTCTCTCTCTCTCTCTCT 7514
 2396 GluGlyGluLeuGlyAspProAspLeuGluProGluGlnValGluProGlnProPro 2415
 7515 GAGGCGGAGCGCGGAGTCCCGATCTC----- 7541
 2416 GlnGlyGlyValAlaAlaProGlySerAspSerGlySerTrpSerThrCysSerGluGlu 2435
 7542 -----AGCGAGCGGTCTTGTGTACCGTAAAGCGGAGGAG 7574
 2436 ---AspAspSerValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuIleThr 2454

Db	3270	TCTGATATGGACCAAGGTTATACCTGGGGGCGAGACACCGCGCGTGGGGACATC	3329	Db	4350	GAGACGGCTGGAGCGCGACTCGTGTGTCTGCCACCGCTACGCGCTCCGGGATCGGTACCC	4409
Qy	1001	LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp	1020	Qy	1361	ThrProHisProAsnIleGluGluValAlaLeuGlyGlnGluGlyGluLeuProPheTyr	1380
Db	3330	ATCTTGGGCTGCGCGCTCTCCGCGCGAGGGAGGAGATACATCTGGGACCGGAGAC	3389	Db	4410	GTGCCACATCCAAACATCAGAGAGTGCTCTGTCCAGACACTGGAGAAATCCCTTTAT	4469
Qy	1021	GlyTyrThrSerLysGlyTyrSerLeuLeuAlaProIleThrAlaTyrAlaGlnGlnThr	1040	Qy	1381	GlyArgAlaIleProLeuSerTyrIleLysGlyArgHisLeuIlePheCysHisSer	1400
Db	3390	AGCCTTGAAGGGCAGGGGTGGCGACTCTTCGCGCTTATACGGCTTCTCCACAGAGC	3449	Db	4470	GGCAAAAGCATCCCATCGAGACCATCAAGGGGGGAGGACCTCATTTCTGCCATCC	4529
Qy	1041	ArgGlyLeuLeuGlyThrIleValSerMetThrGlyArgAspLysThrGluGlnAla	1060	Qy	1401	LysLysLysCysAspGluLeuAlaAlaLeuArgGlyMetGlyLeuAsnSerValAla	1420
Db	3450	CGAGGCTACTTGGCTGCATCATCTAGCTCAGGCGCGGACAGGAACAGGTCGAG	3509	Db	4530	AAGAAGAATGTGATGAGCTCGCCGAGCTGTCCGGCTCGGACTCAATCTGTAGCA	4589
Qy	1061	GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly	1080	Qy	1421	TyTyrArgGlyLeuAspValSerValIleProThrGlnGlnGlyAspValValValAla	1440
Db	3510	GGGGAGGTCCAAAGTGTCTCCACCGCAACACATCTTCTGGCGACCTGGCTCAATGC	3569	Db	4590	TATTACCGGGCTTGTATGATCTCCGTATACCACTAGCGGAGAGCTCATCTCGTAGCA	4649
Qy	1081	ValLeuTyrThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro	1100	Qy	1441	ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnVal	1460
Db	3570	GTGTGTGGACTGTCTATCATGTGTGGCTCAAGACCTTGTCCGCGCCCAAGGGCCCA	3629	Db	4650	ACGGAGCTCTAATCAGCGGCTTTACCGCGATTTCCAGCTCAGTGCATCGACTGCAATACA	4709
Qy	1101	ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTyrProSerProProGly	1120	Qy	1461	AlaValThrGlnValValAspPheSerLeuAspProThrPheThrIleThrGlnIle	1480
Db	3630	ATCACCAATGTACCAATGTGGACAGGACCTCTCGGCTGGCAGCGCCCCCGGG	3689	Db	4710	TGTGTCAACCCAGACAGTGCAGCTTCAGCTTGGACCGGACCTTCACCATTTGAGACGACGACC	4769
Qy	1121	ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrIleValThrArgAsn	1140	Qy	1481	ValProGlnAspAlaValSerArgSerGlnArgArgGlyArgThrGlyArgGlyArgLeu	1500
Db	3690	GCGGTTCTCTTGACACCATCGCTCGGCGAGCTCGGACCTTTTACTTGTGTACAGGCGAT	3749	Db	4770	GTGCCACAAGACCGGCTGTACGCTCGCAGCGCGGAGGAGGAGCTGTGTAGGGGAGGATG	4829
Qy	1141	AlaAspValIleProAlaArgArgArgGlyAspLysArgGlyAlaLeuLeuSerProArg	1160	Qy	1501	GlyIleTyrArgTyrValSerThrGlyGluArgAlaSerGlyMetPheAspSerValVal	1520
Db	3750	GCGGATGTCTTCGCGTCCGCGCGGCGGCGAGCAGCAGGGGGAGGCTTCTCTCCCGAGG	3809	Db	4830	GGCATTTACAGGTTTGTGATCCAGGAGAAACCGCTCGGCGATGTTTCGATTCTCCGTT	4889
Qy	1161	ProLeuSerThrLeuLysGlySerSerGlyGlyProValLeuCysProArgGlyHisAla	1180	Qy	1521	LeuCysGluCysTyrAspAlaGlyAlaAlaTyrTyrGluLeuThrProSerGluThrThr	1540
Db	3810	CCGCTCTCTTCTTGAAGGCTCTTCGGCGGCTCCACTGCTCTGCCCTCGGGGACGCT	3869	Db	4890	CTGTGCGAGTGTATGACGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	4949
Qy	1181	ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle	1200	Qy	1541	ValArgLeuArgAlaTyrPheAsnThrProGlyLeuProValCysGlnAspHisLeuGlu	1560
Db	3870	GTGGGCTCTTCTGGGCTGCGGTGTGACCGGAGGGGTTCGAGGCGGTGGACTTTGTA	3929	Db	4950	GTTAGTTCGGGCTTACCTAAACACACACCGGCTGCGGCTGTGCGCAGACCATCTGGAG	5009
Qy	1201	ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro	1220	Qy	1561	PheTyrGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThr	1580
Db	3930	CCGCTCAGTCTATGGAACACCATATCGGCTCCCGGTCTTACGAGAACACTCGTCCCT	3989	Db	5010	TTCTGGGAGAGCGCTTTTACAGGCTTACCCACCATAGACGCGCATTTCTTGTCCAGACT	5069
Qy	1221	ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLys	1240	Qy	1581	LysGlnSerGlyGluAsnPheAlaTyrLeuThrAlaTyrGlnAlaThrValCysAlaArg	1600
Db	3990	CGGCGGTACCGGACATTCAGGTGGCCCATCTACACGCGCTTACTGTGTAGCGCAAG	4049	Db	5070	AAGCAGCAGGAGACAACTTCCCTTACCTGGTAGCATACAGGCTACGGTGTGCGCAGG	5129
Qy	1241	SerThrLysValProValAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnPro	1260	Qy	1601	AlaLysAlaProProProSerTyrAspValMetTyrLysCysLeuThrArgLeuLysPro	1620
Db	4050	AGCACTAAGGTGGCGCTGCGGTATGACAGCCAAAGGTATAGGTGTCTTCTGAAACCCG	4109	Db	5130	GCTCAGGCTCCACTCCATCGTGGGACCAAAATGTGGAAGTGTCTCATCGGCTAAAGCT	5189
Qy	1261	SerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerLysAlaHisGlyIleAsnPro	1280	Qy	1621	ThrLeuValGlyProThrProLeuLeuTyrArgLeuGlySerValThrAsnGluValThr	1640
Db	4110	TCCGTCGCCGCCACCTTAGTTCGGGCGGTATATGTCTAAGGCGACATGATGACCT	4169	Db	5190	ACGCTCACGGGCCAACCGCTCTGTGTATAGCTGGAGCGCTTCAAACAGGTTACT	5249
Qy	1281	AsnIleArgThrGlyValArgThrValThrGlyAlaProIleThrTyrSerThrTyr	1300	Qy	1641	LeuThrHisProValThrLysTyrIleAlaThrCysMetGlnAlaAspLeuGluValMet	1660
Db	4170	AACATCAGAACCGGGTAGGACCATCACACGGGTGCGCCCATCATCGTACTCCACCTAT	4229	Db	5250	ACCACACACCCCAACCAATACATCATGGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGT	5309
Qy	1301	GlyLysPheLeuAlaAspGlyGlyCysAlaGlyAlaTyrAspIleIleCysAsp	1320	Qy	1661	ThrSerThrTyrValLeuAlaGlyValLeuAlaAlaValAlaAlaTyrCysLeuAla	1680
Db	4230	GGCAAGTTTCTTGGCGAGGTGTGTCTTGGGGCGGCTATGACATCATATATGTAT	4289	Db	5310	ACGAGCACCTTGGGT	5369
Qy	1321	GluCysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGlnAla	1340	Qy	1681	ThrGlyCysValCysIleIleGlyArgLeuHisIleAsnGlnAlaValValAlaPro	1700
Db	4290	GAGTGCCACCTCACTGACTGCACCATCTTCTGGGCGATCGGACAGTCTGGACCAAGCG	4349	Db	5370	ACAGGACGCTGT	5429
Qy	1341	GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProProGlySerValThr	1360	Qy	1701	AspLysGluValLeuTyrGluAlaPheAspGluMetGluCysAlaSerArgAlaAla	1720
Db				Db	5430	GACAGGAGTCTCTTTTACCGGAGTTCGATGATGGAAGAGTGTGTGTGTGTGTGTGTGT	5489

```
QY 261 HisIleAspMetValValMetSerAlaThrLeuCysSerAlaLeuTyrValGlyAspLeu 280
DB 1122 CATGTCGATTTGCTGCTGGGGGGCTGCTCTCTGCTCGCTATGTACGTGGGAGATCTC 1181
QY 281 CysGlyGlyValMetLeuAlaAlaGlnMetPheIleValSerProGlnHisIleHisTrpPhe 300
DB 1182 TGGGAGTCTGTTTCTCGTGGCCAGCTGTTTCACTCTCTCGCTCGCGGACGAGACA 1241
QY 301 ValGlnAspCysAsnCysSerIleTyrProGlyThrIleThrGlyHisArgMetAlaTrp 320
DB 1242 GTACAGGAGTGCATTAATATATATCCCGGACGCTACAGAGTCACTGATGGCTTGG 1301
QY 321 AspMetMetMetAsnTrpSerProThrAlaThrMetIleLeuAlaTyrAlaMetArgVal 340
DB 1302 GATATGATGATGAATGCTCACTACAGAGCCCTAGTGGTATCGCAGTTACTCCGATC 1361
QY 341 ProGluValIleIleAspIleIleSerGlyAlaHisTrpGlyValMetPheGlyLeuAla 360
DB 1362 CCACAAGCTGCTGCTGATATGTTGGGGGGCCCATTTGGGAGTCTTAGCGGCGCTTGGC 1421
QY 361 TyrPheSerMetGlnGlyAlaTrpAlaLysValValIleLeuLeuLeuAlaGly 380
DB 1422 TACTATTCATGTTGGGGAATCGGCTAAGGTTCTGATTTGATGCTACTCTTTTGGCGG 1481
QY 381 ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr 400
DB 1482 GTTGACGGGGACCTATGTACAGGGGGGAGCATGGCCAAACACCCCTCGGGATTACG 1541
QY 401 SerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTrp 420
DB 1542 TCCTCTTTTACCCCGGTCATCCAGAAATCCAGCTTTGTAACACCAACGCGAGCTGG 1601
QY 421 HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer 440
DB 1602 CACATCAACAGGACTCCCTGAATCAATGACTCCCTCAACACTGGGTTCTTGTCTGGC 1661
QY 441 LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg 460
DB 1662 CTGTTCTAGTGACAAAGTTCAACTCATCTGGATGCCAGAGGCGATGCGCAGCTGCAGC 1721
QY 461 SerIleGluAlaPheArgValGlyTrpGlyAlaLeuGlnTyrGluAspAsnValThrAsn 480
DB 1722 CCATCGACGCTTCCTCAGGGGTGGGGCCCATCACTTAC-----AATGAGTCACAC 1775
QY 481 ProGluAspMetArgProTyrCysTrpHisTyrProProArgGlnCysGlyValValSer 500
DB 1776 AGCTCGACACGAGGCTTATTGTTGGCACTACGACCCCGCGCTGCGGTATCTGATACC 1835
QY 501 AlaLysThrValCysGlyProValTyrCysPheThrProSerProValValValGlyThr 520
DB 1836 GGGGGCAGGTGTGTGTCAGTGTACTGCTTACCCCAAGCCCTGTGCTGTGGGGAGC 1895
QY 521 ThrAspArgLeuGlyAlaProThrTyrTrpTrpGlyGluAsnGluThrAspValPheLeu 540
DB 1896 ACCGACCGTTCGGCTCCTACGTACAGTTGGGGGAGATGAGAGCGACGTGTGCTT 1955
QY 541 LeuAsnSerThrArgProProLeuGlySerTrpPheGlyCysThrTrpMetAsnSerSer 560
DB 1956 CTTAAACAACACGCGCGCGCAAGCAACTGTTTGGTGTGTACATGATGAATGACACT 2015
QY 561 GlyTyrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer 580
DB 2016 GGGTTCACCAAGACGTGGCGGGGCCCCCGGTGAACATCGGGGGGATCGGCAATAAACCC 2075
QY 581 ThrAspLeuLeuCysProThrAspCysPheArgLysHisProAspThrThrTyrLeuLys 600
DB 2076 -----TTGACCTGCCCCACGAGCTGCTTCCGGAAGCACCCCGAGGCCACTTACCCAG 2129
QY 601 CysGlySerGlyProTrpLeuThrProArgCysLeuIleAspTyrProTyrArgLeuTrp 620
DB 2130 TGTGGTTCGGGGCTTGGTTGACACCCAGATGTTGGTCTCACTACCCATACAGGCTTGG 2189
```

QY 2975 LysLeuLeuLeuThrProLeuProGluAlaArgLeuLeuAspLeuSerSerTrpPheThr 2994
 Db 9195 AAGCTCAAACTCACTCAATCCCGCTGCGTCCAGTGGATTATCCAGCTGGTTGTT 9254
 QY 2995 ValGlyAlaGlyGlyAspIleThrHisSerValSerArgAlaArgProArgLeuLeu 3014
 Db 9255 GCTGGTTACAGCGGGGAGACATATATCAGCCTGCTGTCGCCGACCGCTGGTTC 9314
 QY 3015 LeuPheGlyLeuLeuLeuPheValGlyValGlyLeuPheLeuLeuProAlaArg 3033
 Db 9315 ATGTGGTGCCTACTCTACTTCTGTAGGGGTAGGACATATCTACTCTCCCAACCGA 9371

RESULT 14
 ABK91429
 ID ABK91429 standard; DNA; 9605 BP.
 XX
 AC ABK91429;
 XX
 DT 15-NOV-2002 (first entry)
 XX
 DE Hepatitis C virus Con 1 isolate DNA mutant 6.
 XX

XX HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;
 XX hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
 XX internal ribosome entry site; IRES; NS5A; HCV replication; mutant.
 XX
 OS Hepatitis C virus.
 OS Synthetic.

XX Key Location/Qualifiers
 FH 342..9374
 FT /*tag= a
 FT /product= "HCV polyprotein"
 FT /note= "The polyprotein consists of the Core, E1,
 FT E2, P1, NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
 FT mutation
 FT replace (6859,C)
 FT /*tag= b

XX W0200259321-A2.
 XX
 XX 01-AUG-2002.
 XX
 XX 16-JAN-2002; 2002W0-EP00526.
 XX
 XX 23-JAN-2001; 2001US-263479P.
 XX
 XX (RICE-) IST RICERHE BIOL MOLECOLARE ANGELETTI.
 XX
 XX De Francesco R, Migliaccio G, Paonessa G;
 XX WPI; 2002-599793/64.
 XX

XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
 PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
 PT ribosome entry site (IRES) region, useful in studying HCV replication
 PT and expression
 XX
 XX Claim 9; Page -: 69pp; English.

XX The invention relates to nucleic acid molecules comprising altered HCV
 CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
 CC internal ribosome entry site (IRES) region coding for one or more NS3,
 CC NS5A, or EMCV IRES mutations, respectively. The location of the
 CC mutations are detailed in the specification. Also included are
 CC (1) an expression vector comprising a nucleotide sequence coding for
 CC the altered nucleic acids, which is transcriptionally coupled to an
 CC exogenous promoter; (2) a recombinant cell human hepatoma cell comprising
 CC the altered nucleic acids; (3) a recombinant cell produced by introducing
 CC into a human hepatoma cell the altered nucleic acids; (4) producing an
 CC HCV (hepatitis C virus) replicon enhanced cell or which containing a
 CC functional HCV replicon; (5) an HCV replicon enhanced cells made in the
 CC method; and (6) measuring the ability of a compound to affect HCV
 CC activity. The HCV replicons and HCV replicon enhanced cells are useful in

CC studying HCV replication and expression, and HCV and host cell
 CC interactions, producing HCV RNA and proteins, and providing a system
 CC for measuring the ability of a compound to modulate one or more HCV
 CC activities e.g. to discover drugs which may treat HCV mediated
 CC diseases such as liver failure, cirrhosis and hepatocellular carcinoma.
 CC The present sequence is an HCV replicon Con 1 mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the HCV sequence appearing as
 CC ABK91411 and the information in Claim 9.
 XX
 XX Sequence 9605 BP; 1910 A; 2882 C; 2733 G; 2080 T; 0 other;

Alignment Scores:
 Pred. No.: 0 Length: 9605
 Score: 12031.50 Matches: 2183
 Percent Similarity: 83.42% Conservative: 352
 Best Local Similarity: 71.83% Mismatches: 469
 Query Match: 74.64% Indels: 35
 DB: 24 Gaps: 7

US-09-980-559-2 (1-3033) x ABK91429 (1-9605)
 QY 1 MetSerThrAsnProGlyProGlnArgLysThrLysArgAsnThrAsnArgArgProGln 20
 Db 342 ATGAGCAGCAATCTCTAAACCTCAAGAAAAACCAACGTAACCAACCGCCGCCACAG 401
 QY 21 AspValLysPheProGlyGlyGlnIleValGlyGlyValLysLeuLeuProArg 40
 Db 402 GAGGTCAAGTTCCTCGGGCGGTGTCAGATCGTCGGTGGAGTTACTGTTCGGCGCAGG 461
 QY 41 GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly 60
 Db 462 GCGCCCGAGTTCGGTGGCGCGACTAGAAAGACTTCCGAGCGGTGCGAACCTCTCGTGA 521
 QY 61 ArgArgGlnProIleProLysAspArgArgSerThrGlyLysSerTrpGlyLysProGly 80
 Db 522 AGCGGACAACTATATCCCAAGGCTCGCCAGGCTCGCCAGCGGTAGGCGCTGAGCCCGG 581
 QY 81 TyrProTrpProLeuTyrGlyAsnGluGlyLeuGlyTrpAlaGlyTrpLeuLeuSerPro 100
 Db 582 TACCCCTGCGCCCTCTATGCAATGAGGCTTGGGTGGCGAGGATGGCTCTCTGTCAACC 641
 QY 101 ArgGlySerArgProSerTrpGlyProAsnAspProArgHisArgSerArgAsnValGly 120
 Db 642 CGTGGCTCTCGGCTAGTTGGGGCGCCCGACGACCCCGCGGTAGGTGCGGCAATTTGGGT 701
 QY 121 LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetClyTyrIleProValVal 140
 Db 702 AAGGTATCATGATACCTCCTCGCGCTTCGCGGATCTCATGGGTGATATTCGCTCGTC 761
 QY 141 GlyAlaProLeuGlyGlyValAlaArgAlaLeuAlaHisGlyValArgValLeuGluAsp 160
 Db 762 GGGCGCCCTTAGGGGGCGCTGCCAGGCGCTCGCGCATGGCGTCCGGGTTCCTGGAGGAC 821
 QY 161 GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla 180
 Db 822 GCGGTGAACATATGCAACAGGGAATCTCGCGGTGCTCTCTCTTCTATCTCTCTTTGGCT 881
 QY 181 LeuLeuSerCysIleThrThrProValSerAlaAlaGluValLysAsnIleSerThrGly 200
 Db 882 TTGCTGTCTGTGTGTGACCATCCAGCTTCGCTTATGAGTGGCAACGATATCCGGAGTG 941
 QY 201 TyrMetValThrAsnAspCysThrAsnAspSerIleThrTrpGlnLeuGlnAlaVal 220
 Db 942 TACCATGTCAAGCACTGCTCCCAACGCAAGCATTTGTGTATGAGGCGGCGGACATGATC 1001
 QY 221 LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTrpIle 240
 Db 1002 ATGCATACCCCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1061
 QY 241 ProValSerProAsnValAlaValGlnArgProGlyAlaLeuThrGlnGlyLeuArgThr 260
 Db 1062 GCGCTCACTCCACGCTCG 1121

2615	MetAlaLeuTyrAspIleThrGlnLysLeuProGlnAlaValMetGlyAlaSerTyrGly	2634
8115	ATGGGCCCTTTACGATGTGGTCTCCACCCCTCCATGGCGGTGATGGGCTCTTCATACGGA	8174
2635	PheGlnTyrSerProAlaGlnArgValGluPheLeuLeuLysAlaTrpAlaGluLysLys	2654
8175	TTCCAAATACTCTCCCTGGACAGCGGTCCAGTTCTCGGTGAATGCTCGAAAGCGAAGAAA	8234
2655	AspProMetGlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluArgAsp	2674
8235	TGCCCTATGGGCTTCGCATATGACACCCCGCTGTTTGTACTCAACGGTCTCAGTAATGAC	8294
2675	IleArgThrGluGluSerIleTyrArgAlaCysSerLeuProGluGluAlaHisThrAla	2694
8295	ATCCCGTGTGGAGAGTCAATCTACCAATGTTGTGTACTTTGGCCCCCGAAGCAGACAGGCC	8354
2695	IleHisSerLeuThrGluArgLeuTyrValGlyGlyProMetPheAsnSerLysGlyGln	2714
8355	ATAAGGTCCCTCACAGAGCGCTTTACATCGGGGGCCCCCTGACTAATCTCTAAAGGGCAG	8414
2715	ThrCysGlyTyrArgArgCysArgAlaSerGlyValLeuThrThrSerMetGlyAsnThr	2734
8415	AACTCGCGGTATCGCGCGTGGCGCGAGCGGTGTACTGACGACCGACGCTGGGTAAATACC	8474
2735	IleThrCysTyrValLysAlaLeuAlaAlaCysLysAlaAlaGlyIleAlaProThr	2754
8475	CTCACATGTTACTTGGAGCGCTGGCGCCTGTGAGCTGCGAAGCTCCAGGACTGCGACG	8534
2755	MetLeuValCysGlyAspAspLeuValIleSerGluSerGlnGlyThrGluGluAsp	2774
8535	ATGCTCGTATCGGAGACGACCTTGCTGTTATCTGTGAAGCGGGGACCCCAAGAGGAC	8594
2775	GluArgAsnLeuArgAlaPheThrGluAlaMetThrArgTyrGlySerAlaProProGlyAsp	2794
8595	GAGGCGAGCCTACGGGCCCTTACGAGAGCTATGACTAGATACTCTGCCCGCCCTCGGGAC	8654
2795	ProProArgProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerVal	2814
8655	CGGCCCAACAGAAATACGACTTGGAGTTGATACATCATGCTCTCCAAATGTGTCACTC	8714
2815	AlaLeuGlyProGlnGlyArgArgTyrTyrIleThrArgAspProThrThrProIle	2834
8715	GCGCACGATGCACTCTGGCAAAAGGTGTACTATCTCACCGCTGACCCACACCCCGCTT	8774
2835	AlaArgAlaAlaTrpGluThrValArgHisSerProValAsnSerTrpLeuGlyAsnIle	2854
8775	GCGCGGGCTGGTGGAGACAGCTAGACACACTCCAGTCAATCTCTGGCTAGGCACATC	8834
2855	IleGlnTyrAlaProThrIleTrpAlaArgMetValLeuMetThrHisPheSerIle	2874
8835	ATCATGTATGGCCCCACCTTGTGGGCAAGGATGATCCTGTGACTCATTTCTTCTCCATC	8894
2875	LeuMetAlaGlnAspThrLeuAspGlnAsnLeuAsnPheGluMetTyrGlyValAlaValTyr	2894
8895	CTTCTAGCTCAGGAAACAATCTGAAAGGCCCTAGATTGTCTAGATCTACGGGGCTGTGTAC	8954
2895	SerValSerProLeuAspLeuProAlaIleIleGluArgLeuHisGlyLeuAspAlaPhe	2914
8955	TCCATTGAGCCACTTGACCTACCTCAGATCATTTCAACGACTCCATGGGCTTTAGCGCAIT	9014
2915	SerLeuHisThrTyrThrProHisGluLeuThrArgValAlaSerAlaLeuArgLysLeu	2934
9015	TCACATCCATAGTACTCTCCAGGTGAGATCAATAGGGTGGCTTATGCTCTCAGGAACCT	9074
2935	GlyAlaProProLeuArgAlaTrpLysSerArgAlaArgAlaValArgAlaSerLeuIle	2954
9075	GGGGTACCGCCCTTGGAGAGCTGGAGACATCGGGCCAGAGTGTCCGGCTAGGCTACTG	9134
2955	SerArgGlyArgAlaAlaValCysGlyArgTyrIleuPheAsnTrpAlaValLysThr	2974
9135	TCCACGGGGGAGCGGCTGCACCTTGTGGCAAGTACTCTTCACTGGGCGAGTAAGGACC	9194

5970	:::	GAGGACTGGTTAACTACTCCCTGCTATCTCTCCCTCGGCCCTAGTCTGCGGGTC	6029
5989	:::	IleCysAlaalaileuArgHisValGlyProGlyGluGlyAlaValGlntrpMet	1920
5990	:::	GTGTGCGCAGCGATACTGCGTCGGCAGCGTGGGCCACGSGAGGGGCTGTGTCAAGTGGATG	6089
1921	ASnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaProThrHisTyrValThr	1940	
6090	AACCGGCTGATAGCGTTGCGTTTCGCGGGGTAAACCACTCTCTCCCCACGCACTATGTGCGCT	6149	
1941	GluSerAspAlaSerGlnArgValThrGlnLeuLeuGlySerLeuThrIleThrSerLeu	1960	
6150	GAGAGCGAGCTGCGACGAGTGTCACTCAGATCTCTCTTAGTCTTACCATCACTCAGCTG	6209	
1961	LeuArgLeuGluHisAsnTrpIleThrGluAspCysProIleProCysGlyGlySerTrp	1980	
6210	CTCAAGAGGCTTACCAGTGGATCAACAGGAGCTGCTCCAGCGCATGCTCCGGCTCGTGG	6269	
1981	LeuArgAspValTrpAspTrpValCysThrIleLeuThrAspPheLysAsnTrpLeuThr	2000	
6270	CTAAGAGATGTTTGGATTGGATATGCACGGTGTGTGATCTGATTTTCAAGACCTCGCTCCAG	6329	
2001	SerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTyrLys	2020	
6330	TCCAAGCTCTGCGCGGATTTGCGGGAGTGCCTCTTCTCATGTCAACGTGGGTACAAG	6389	
2021	GlyValTrpAlaGlyThrGlyIleMetThrThrArgCysProCysGlyAlaAsnIleSer	2040	
6390	GGAGTCTGGCGGGCGACGGCATCATGCAACACCTGCCCATGTGGAGCACAGATCACC	6449	
2041	GlyAsnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTrp	2060	
6450	GGACATGTTGAAAAACGGTTTCCATGAGGATCGTGGGGCTTAGGACCTCTGTAGTAACACGTGG	6509	
2061	GlnGlyThrPheProIleAsnCysTyrThrGluGlyGlnCysValProLysProAlaPro	2080	
6510	CATGGAAACATTTCCCATTAACCGGTACACAGGGGCCCTCGACGCCCTTCCCGCGGCCA	6569	
2081	AsnPheLysValAlaIleTrpArgValAlaAlaSerGluTyrAlaGluValThrGlnHis	2100	
6570	AAATATTCTAGGGCGCTGTGGCGGTGGCTGCTGAGGAGTAGCTGGAGGTTACGCGGGTG	6629	
2101	GlySerTyrHisTyrIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeu	2120	
6630	GGGGATTTCCACTACGTGACGGGCATGACCATGACAACTGCTCCCGTGTGTCAGGTT	6689	
2121	ProSerProGluPhePheSerTrpValAspGlyValGlnIleHisArgPheAlaProThr	2140	
6690	CCGGCCCCGGAATTTCTCACAGAGTGGATGGGGTGGCGTGTGCACAGGTAGCTTCCAGCG	6749	
2141	ProLysProPhePheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal	2160	
6750	TGCNAACCTCTCTACGGGAGAGGTGCATATCTCTGGTTCGGGCTCAATCAATACCTGTT	6809	
2161	GlySerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThr	2180	
6810	GGGTCAAGCTCCCATGCGAGCCGAAACCGGACGTAGCAGTGTCTCACTTCCATGCTCACC	6869	
2181	AspProSerHisIleThrAlaGluThrAlaAlaArgArgLeuAlaArgGlySerProPro	2200	
6870	GACCCCTCCACATTAACGCGAGACGGCTAAGCGTAGGCTGGCCAGGGAGATCTCCCCC	6929	
2201	SerGluAlaSerSerAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThr	2220	
6930	TCCCTGGCCAGCTCATCAGCTAGCCAGCTGTCTGCGCCTCTCTTGAAGGCCAATGCACT	6989	
2221	ThrHisGlyLysAlaTyrAspValAspMetValAspAlaAsnLeuPhe-----	2236	
6990	ACCGTCACTGACTCCCGGACGCTGACCTCATCGAGGCCAACCTCTGCTGGCGCAGGAG	7049	
2237	MetGlyGlyAspValThrArgIleGluSerGlySerLysValValValLeuAspSerLeu	2256	

Db 2670 TACATCAAGGCGAGGCTGGTCCCTGGGGGGGCATATGCCCCCTCTACGCGCTATGGCCGCTA 2729
Qy 801 SerLeuLeuLeuLeuAlaLeuProGlnGlnAlaTyrAlaTyrAspAlaSerValHisGly 820
Db 2730 CTCCTCTCTGCTGGCGTTACCAACAGGAGATACGCCATGAGCGGAGATGGCAGCA 2789
Qy 821 GlnIleGlyAlaAlaLeuLeuValMetIleThrLeuPheThrLeuThrProGlyTyrLys 840
Db 2790 TCGTGGGAGGCGGCTTTTCGTAGTCTGATCTCTTGACCTTGTACCGGCACTATAAG 2849
Qy 841 ThrLeuLeuSerArgPheLeuTrpTrpLeuCysTyrLeuLeuThrLeuGlyGluAlaMet 860
Db 2850 CTGCTCTCGCTAGGCTCATAGTGTGTACATATTTATACCAAGGCGGAGGCACAC 2909
Qy 861 ValGlnGluTrpAlaProMetGlnValArgGlyGlyArgAspGlyIleIleTrpAla 880
Db 2910 TTGCAAGTGTGATCCCTCCCTCAACGTTTGGGGGGCGGATGCCGTCTACCTCTCCTC 2969
Qy 881 ValAlaIlePheTyrProGlyValValPheAspIleThrLysTrpLeuLeuAlaValLeu 900
Db 2970 ACGTGGCGATCCACCCAGAGCTAATCTTTACCATCACCAAAATCTTGCTCGCCATCTC 3029
Qy 901 GlyProAlaTyrLeuLeuLysGlyAlaLeuThrArgValProTyrPheValArgAlaHis 920
Db 3030 GGTCCACTCATGTGCTCCAGGCTGTATAAACAAGTCCGTACTTCTGTGGCGGCACAC 3089
Qy 921 AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyGlyArgTyrValGlnMet 940
Db 3090 GGGCTCATTCGTGCATGCTGCTGGTGGGAAGTTGCTGGGGGTCTATTATGTCCAAATG 3149
Qy 941 AlaLeuLeuAlaLeuGlyArgTrpThrGlyThrTyrIleTyrAspHisLeuThrProMet 960
Db 3150 GCTCTCATGAGTTGGCCGCACTGACGAGTACGTAGTTTATGACCATCTCACCCACATG 3209
Qy 961 SerAspTrpAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIleIlePhe 980
Db 3210 CGGGACTGGGGCCACCGGGCTTACGAGACCTTGGCGTGGCAGTTGAGCCGCTGCTTC 3269
Qy 981 SerProMetGluLysIleValIleValTrpGlyAlaGluThrAlaAlaCysGlyAspIle 1000
Db 3270 TCTGATATGGAGACCAAGGTTATCACCTGGGGGGCAGACACCGCGCGTGTGGGGACATC 3329
Qy 1001 LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp 1020
Db 3330 ATCTGGGCTGCTGCTCCGCCCGAGGGGAGGAGATACATCTGGGACCGGACAC 3389
Qy 1021 GlyTyrThrSerLysGlyTrpSerLeuLeuAlaProIleThrAlaTyrAlaGlnThr 1040
Db 3390 AGCCTTGAAGGGCAGGGGTGGCGACTCTCTCGGCTTATTACGGCTTACTCCCAACAGACG 3449
Qy 1041 ArgGlyLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGluGlnAla 1060
Db 3450 CGAGGCTTACTTGGCTGCATCATCTAGCCCTCACAGGCGGGAGAGAACCGAGTGCAG 3509
Qy 1061 GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly 1080
Db 3510 GGGAGGTCAGTGTCTCCACCGCAACACATCTTCTGGCAGCTTGGCTCATGTC 3569
Qy 1081 ValLeuTrpThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro 1100
Db 3570 GTGTGTGGACTGTCTATCATGTGTCGGCTCAAGACCTTGGCGGCCCAAGGGGCCCA 3629
Qy 1101 ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTrpProSerProGly 1120
Db 3630 ATCACCCAAATGTACCAATGTGGACACGAGCTCTGCTGGTGGCAAGCGCCCCCGGG 3689
Qy 1121 ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuThrLeuValThrArgAsn 1140
Db 3690 GCGGTTCTTACACCATGACCTCGCGCAGCTTACCTTACCTTGTGTACAGGCAAT 3749
Qy 1141 AlaAspValIleProAlaArgArgGlyAspLysArgGlyAlaLeuLeuSerProArg 1160
Db 3750 GCCGATGTCTTCGGTGGCGGGCGGCGGCGAGCAGCAGGGGGGAGCTTCTCTCCCGCAGG 3809

Qy 1161 ProLeuSerThrLeuLysGlySerSerGlyGlyProValLeuCysProArgGlyHisAla 1180
Db 3810 CCGCTCTCTACTTGAAGGGCTCTTCGGCGGTTCACCTCTCTGCCCTCGGGGACGCT 3869
Qy 1181 ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle 1200
Db 3870 GTGGGCATCTTTCGGCTGCGGTGACCCGAGGGTTCGAAAGCGGTGGACTTTGTA 3929
Qy 1201 ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro 1220
Db 3930 CCGCTCGAGTCTATGGAAACCACTATGCGGTCCCGGTCTTCAAGCAACTCGTCCCTC 3989
Qy 1221 ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLys 1240
Db 3990 CCGGCGGTACCGACACATTCAGGTGGCCATCTACAGCCCTCTACTGGTAGGCGAAG 4049
Qy 1241 SerThrLysValProValAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnPro 1260
Db 4050 AGCACTAAGTCCCGCTGCGTATGCAGCCCAAGGTATAAGGTGCTTCTCTGAACCCG 4109
Qy 1261 SerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerLysAlaHisGlyIleAsnPro 1280
Db 4110 TCCGTGCGCGCCACCTAGTTCGGGGGTATATGTCTAAGGCACATGGTATCGACCTC 4169
Qy 1281 AsnIleArgThrGlyValArgThrValThrThrGlyAlaProIleThrTyrSerThrTyr 1300
Db 4170 AACATCAGAACCGGGGTAAAGACCATCACCGGTGCCCCCATCACGTACTCCACTAT 4229
Qy 1301 GlyLysPheLeuAlaAspGlyGlyCysAlaGlyAlaTyrAspIleIleIleCysAsp 1320
Db 4230 GCCAAGTTCTTGGCGAGCGTGTGCTCTGGGGCGCTATGACATCATATATGTAT 4289
Qy 1321 GluCysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGlnAla 1340
Db 4290 GAGTGCACATCACTGACTCGACCATCTCTCGGCATCGGCACAGTCTCGACCAAGCG 4349
Qy 1341 GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProGlySerValThr 1360
Db 4350 GAGACGGCTGGAGCGGACTCTGCTGCTCGCCACCGCTACGCTCCGGGATCGGTACC 4409
Qy 1361 ThrProHisProAsnIleGluValAlaLeuGlyGlnGluGlyIleProPheTyr 1380
Db 4410 GTGCCACATCCAAACATCGAGGAGTGGCTCTGTCCAGCACTGGAGAAATCCCTTTAT 4469
Qy 1381 GlyArgAlaIleProLeuSerTyrIleLysGlyArgHisLeuIlePheCysHisSer 1400
Db 4470 GGCAAAGCCATCCCATCGAGACCATCAAGGGGGGAGGCACCTCATTTTCTGCCATTC 4529
Qy 1401 LysLysLysCysAspGluLeuAlaAlaLeuArgGlyMetGlyLeuAsnSerValAla 1420
Db 4530 AAGAAGAATGTATGATGCTCCCGGAAGCTGTCCGGCTCGGACTCAATGCTGTAGCA 4589
Qy 1421 TyrTyrArgGlyLeuAspValSerValIleProThrGlnGlyAspValValValAla 1440
Db 4590 TATTACCGGGGCTTGTATGATATCCGTATACCACTAGCGGAGAGCTCATTTGTGTAGCA 4649
Qy 1441 ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnVal 1460
Db 4650 ACGGACGCTTAATACGGGCTTTTACCGGCAATTCAGTCTAGTATCGACTGCAATACA 4709
Qy 1461 AlaValThrGlnValValAspPheSerLeuAspProThrPheThrIleThrGlnIle 1480
Db 4710 TGTGTACCCAGACAGTTCAGCTTCAGCTTGGACCCGACCTTCAACATGAGACGAGACC 4769
Qy 1481 ValProGlnAspAlaValSerArgSerGlnArgArgGlyArgThrGlyArgGlyArgLeu 1500
Db 4770 GTGCCACAAGACCGGCTGTACGCTCGCAGCGCGGAGGAGGAGTGTGTAGGGGACGATG 4829
Qy 1501 GlyIleTyrArgTyrValSerThrGlyGluArgAlaSerGlyMetPheAspSerValVal 1520
Db 4830 GGCATTTACAGTGTGTGACTCCAGGAGAACCGCCCTCGGGCATGTGTGATTTCTCGGTT 4889


```

Db      8535 ATCTCGTATGCGGAGACGACCTTGTCGTTATCTGTGAAGCGCGGGACCCCAAGAGGAC 8594
Qy      2775 GluArgAsnLeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAsp 2794
Db      8595 GAGCGAGCTACGGCGCTTCACGGAGCTATGACTAGATACCTCTGCCCGCCCTGGGGAC 8654
Qy      2795 ProProArgProGluTyrAspLeuGluLeuLeuThrSerCysSerSerAsnValSerVal 2814
Db      8655 CCGCCCAACACAGATACGATTTGGAGTTGATTAACATCATGCTCCTCCAATGTGTGATC 8714
Qy      2815 AlaLeuGlyProGlnGlyArgArgTyrTyrLeuThrArgAspProThrProIle 2834
Db      8715 GCGCAGCATCATCTGCAAAAGGCTGATCTATCTCACCGTGACCCACCCGCCCT 8774
Qy      2835 AlaArgAlaAlaTrpGluThrValArgHisSerProValAsnSerTrpLeuGlyAsnIle 2854
Db      8775 GCGCGGCTGCGTGGGAGACAGCTAGACACATCCAGTCAATTCCTGGGTAGCAACATC 8834
Qy      2855 IleGlnTyrAlaProThrIleTrpAlaArgMetValLeuMetThrHisPheSerIle 2874
Db      8835 ATCATGTATGCGCCACCTTGTTGGCAAGATCATCTGATGACTCATTTCTTCTCCATC 8894
Qy      2875 LeuMetAlaGlnAspThrLeuAspGlnAsnLeuAsnPheGluMetTyrGlyAlaValTyr 2894
Db      8895 CTCTAGCTCAGGAACAACCTGGAAGAGCCCTAGATTGTACATCTACGGGCGCTGTAC 8954
Qy      2895 SerValSerProLeuAspLeuProAlaIleIleGluArgLeuHisGlyLeuAspAlaPhe 2914
Db      8955 TCATTGAGCCATTTGACCTACCTTCAGATCATTTCAACGACTCCATGGCCTTAGCGCATTT 9014
Qy      2915 SerLeuHisThrTyrThrProHisGluLeuThrArgValAlaSerAlaLeuArgLysLeu 2934
Db      9015 TCATCTCAATGTTACTCTCCAGTGAGATCAATAGGTGGTTCATGCTCAGGAACATT 9074
Qy      2935 GlyAlaProProLeuArgAlaTrpLysSerArgAlaArgAlaValArgAlaSerLeuIle 2954
Db      9075 GGGGTACCGCCCTTGGCGAGTCTGGAGACATCGGCCAGAAAGTGTCCGCGTAGGCTACTG 9134
Qy      2955 SerArgGlyGlyArgAlaAlaValCysGlyArgTyrLeuPheAsnTrpAlaValLysThr 2974
Db      9135 TCCAGGGGGGAGGCTGCCACTTGTGCAAGTACTCTTCAACTGGGCGAGTAAGGACC 9194
Qy      2975 LysLeuLysLeuThrProLeuProGluAlaArgLeuLeuAspLeuSerSerTrpPheThr 2994
Db      9195 AGCTCAAACTCACTCCATCCGCTGGTCCAGTTGGATTATCCAGCTGGTTCGTT 9254
Qy      2995 ValGlyAlaGlyGlyAspIleTyrHisSerValSerArgAlaArgProArgLeuLeu 3014
Db      9255 GCTGTGTACAGCGGGGAGACATATATACAGCTGTCTCGTCCCGACCCCGCTGGTTC 9314
Qy      3015 LeuPheGlyLeuLeuLeuPheValGlyValGlyValGlyLeuPheLeuLeuProAlaArg 3033
Db      9315 ATGTGGTGCTACTCTCTACTTTCTGTAGGGTAGGCATCTATCTACTCCCAACCGA 9371

RESULT 13
ID ABK91411
XX ABK91411 standard; DNA; 9605 BP.
AC ABK91411;
XX
DT 15-NOV-2002 (first entry)
DE Hepatitis C virus Con 1 isolate DNA.
XX
KW HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;
KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
KW internal ribosome entry site; IRES; NS5A; HCV replication.
XX
OS Hepatitis C virus.
XX
XX Key Location/Qualifiers
FH 342..9374
FT CDS

```

```

FT      1/*tag= a
FT      2/product= "HCV polyprotein"
FT      3/Note= "The polyprotein consists of the Core, E1,
XX      4E2, P7, NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
PN      5WO200259321-A2.
XX
XX      601-AUG-2002.
XX
XX      16-JAN-2002; 2002WO-EP00526.
XX
XX      23-JAN-2001; 2001US-363479P.
XX
XX      (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
XX
XX      De Francesco R, Migliaccio G, Paonessa G;
XX      WPI; 2002-599793/64.
XX      P-PSDB; ABG32451.
XX
XX      New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
XX      NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
XX      ribosome entry site (IRES) region, useful in studying HCV replication
XX      and expression
XX
XX      Claim 9; Page 36-39; 69pp; English.
XX
XX      The invention relates to nucleic acid molecules comprising altered HCV
XX      NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
XX      internal ribosome entry site (IRES) region coding for one or more NS3,
XX      NS5A, or EMCV IRES mutations, respectively. The location of the
XX      mutations are detailed in the specification. Also included are
XX      (1) an expression vector comprising a nucleotide sequence coding for
XX      the altered nucleic acids, which is transcriptionally coupled to an
XX      exogenous promoter; (2) a recombinant cell human hepatoma cell comprising
XX      the altered nucleic acids; (3) a recombinant cell produced by introducing
XX      into a human hepatoma cell the altered nucleic acids; (4) producing an
XX      HCV (hepatitis C virus) replicon enhanced cell or which containing a
XX      functional HCV replicon; (5) an HCV replicon enhanced cells made in the
XX      method; and (6) measuring the ability of a compound to affect HCV
XX      activity. The HCV replicons and HCV replicon enhanced cells are useful in
XX      studying HCV replication and expression, and HCV and host cell
XX      interactions, producing HCV RNA and proteins, and providing a system
XX      for measuring the ability of a compound to modulate one or more HCV
XX      activities e.g. to discover drugs which may treat HCV mediated
XX      diseases such as liver failure, cirrhosis and hepatocellular carcinoma.
XX      The present sequence is the HCV replicon Con 1, used as a basis for
XX      the adaptive mutations of the invention.
XX
XX      Sequence 9605 BP; 1910 A; 2883 C; 2733 G; 2079 T; 0 other;

```

```

Alignment Scores:
Pred. No.: 0 Length: 9605
Score: 12031.50 Matches: 2183
Percent Similarity: 83.42% Conservative: 352
Best Local Similarity: 71.83% Mismatches: 469
Query Match: 74.64% Indels: 35
DB: 24 Gaps: 7

US-09-980-559-2 (1-3033) x ABK91411 (1-9605)
Qy      1 MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgProGln 20
Db      342 ATGACGACGATCTCTAACTTCAAGAAACCAACCAAGTATACACCAACGCGCCGACAG 401
Qy      21 AspValLysPheProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArg 40
Db      402 GAGCTCAAGTTCCCGCGCGGTGGTTCAGATCGTCGGTGGAGTTTACCTGTTCCGCGCAGG 461
Qy      41 GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly 60
Db      462 GCGCCCAAGTTGGGTGTGCGCGAGTTCGAGAGACTTCGAGCGGTTCGCAACCTCGTGA 521

```

Qy	2041	GlyAsnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTyr	2060	Qy	2396	GluGlyGluLeuGlyAspProAspLeuGluProGluGlnValGluProGlnProProPro	2415
Db	6450	GGACGTGTGAAACCGGTTTCATAGAGATCTGGGCTTAGGACCTGTAGTAAACAGTGG	6509	Db	7515	GAGGGGAGCGGGGATCCCCGATCTC	7541
Qy	2061	GlnGlyThrPheProIleAsnCysTyrThrGluGlyGlnCysValProLysProAlaPro	2080	Qy	2416	GlnGlyGlyValAlaAlaProGlySerAspSerGlySerTyrThrCysSerGluGlu	2435
Db	6510	CATGGAACATTCCTCCATTAACGGGTACACACCGGGCCCTGCACGGCCCTCCCGGGGCCA	6569	Db	7542	-----ACGGACGGGTCTTGCTGTACCGTAAGCGAGGAG	7574
Qy	2081	AsnPhelysValAlaIleTyrArgValAlaAlaSerGluTyrAlaGluValThrGlnHis	2100	Qy	2436	---AspAspSerValValCysCysSerMetSerTyrThrTyrThrGlyAlaLeuIleThr	2454
Db	6570	AAATATTCTTAGGGCGCTGTGGGGTGGCTGCTGAGAGTACGTGAGGTTACGGGGTG	6629	Db	7575	GCTAGTGAGGAGCTGCTGCTGCTGCTACATGCTACATGACAGGCGCCCTGATCAG	7634
Qy	2101	GlySerTyrHisTyrIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeu	2120	Qy	2455	ProCysSerProGluGluGluLysLeuProIleAsnProLeuSerAsnSerLeuLeuArg	2474
Db	6630	GGGGATTTCCATCTAGCGGAGCATGACCACTGACAACTGAAGTCCCGTGTCTCAGTT	6689	Db	7635	CCATGCGCTGCGGAGGAAACCAAGCTGCCATCAATGCACTGAGCAACTCTTTGTCTCCGT	7694
Qy	2121	ProSerProGluPhePheSerTyrValAspGlyValGlnIleHisArgPheAlaProThr	2140	Qy	2475	TyrHisAsnLysValTyrCysThrThrThrLysSerAlaSerLeuArgAlaLysVal	2494
Db	6690	CCGGCCCCGAAATTTCTTACAGAAAGTGGATGGGTGGCTTGCACAGGTACGCTCCAGCG	6749	Db	7695	CACCAACAATCTGCTATGCTACCACTCTCCGACGCGCAAGCTGCGGAGGAAAGGTC	7754
Qy	2141	ProLysProPhePheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal	2160	Qy	2495	ThrPheAspArgMetGlnValLeuAspSerTyrTyrAspSerValLeuLysAspIleLys	2514
Db	6750	TGCAAAACCCCTCTACCGGAGGAGGTACATTTCTGCTGCGGCTCAATCAATACCTGGTT	6809	Db	7755	ACCTTTGACAGACTGAGGTCTTGACGACCACTACCGGAGCTGCTCAAGGAGATGAAG	7814
Qy	2161	GlySerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThr	2180	Qy	2515	LeuAlaSerLysValThrAlaArgLeuLeuThrMetGluGluAlaCysGlnLeuThr	2534
Db	6810	GGGTACAGCTCCCATCGAGCCCGAACCGAGCTAGCAGTGTCTCATCTTCATGCTCACC	6869	Db	7815	GCGMAGCGCTCCACAGTTTAAAGCTTAACTTCTATCGTGGAGGAGCTGTAAGCTGACG	7874
Qy	2181	AspProSerHisIleThrAlaGluThrAlaAlaArgArgLeuAlaArgGlySerProPro	2200	Qy	2535	ProProHisSerAlaArgSerLysTyrGlyPheGlyAlaLysGluValArgSerLeuSer	2554
Db	6870	GACCCCTCCACATTTACGGCGAGACGGCTAAGCGTAGGCTGGCGACGGGATCTCCCCC	6929	Db	7875	CCCCCACAATTCGGCCAGATCTAAATTTGGCTATGGGGCAAGAGAGCTCCGGAACCTTATCC	7934
Qy	2201	SerGluAlaSerSerAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThr	2220	Qy	2555	GlyArgAlaValAsnHisIleLysSerValTyrLysAspLeuLeuGluAspSerGluThr	2574
Db	6930	TCCTTGCCAGCTCATCAGCTAGCAGCTGTCTGCGCTTCTTGAAGGCAACATGCAC	6989	Db	7935	AGCAAGCGCTTAACCAATCCGCTCCGCTGGAAGGACTTCTGGAAGACACTGAGACA	7994
Qy	2221	ThrHisGlyLysAlaTyrAspValAspMetValAspAlaAsnLeuPhe	2236	Qy	2575	ProIleProThrThrIleMetAlaLysAsnGluValPheCysValAspProThrLysGly	2594
Db	6990	ACCGCTATGACTCCCGGACCGCTGACTCATCGAGCCCACTCTCTGTCGGCGCAGGAG	7049	Db	7995	CCAAATTTGACACCACCATCATGCAAAATAGGTTTTCTGCTCCCAACAGAGAGGGG	8054
Qy	2237	MetGlyGlyAspValThrArgIleGluSerGlySerLysValValValLeuAspSerLeu	2256	Qy	2595	GlyLysLysAlaAlaArgLeuIleValTyrProAspLeuGlyValArgValCysGluLys	2614
Db	7050	ATGGCGGGAACATCACCGCGTGGAGTGCAGAAATAAGGTAGTAAATTTGGACTCTTC	7109	Db	8055	GGCGCAAGCCAGCTCGCTTATCGTATTTCCAGATTTCCGAGTTTGGGGTTCGTGTGCGAGAA	8114
Qy	2257	AspProMetValGluLysArgSerAspLeuGluProSerIleProSerGluTyrMetLeu	2276	Qy	2615	MetAlaLeuTyrAspIleThrGlnLysLeuProGlnAlaValMetGlyAlaSerTyrGly	2634
Db	7110	GAGCGCTCCAAAGCGGAGGAGATGAGAGGGAAGTATCCGTTCGCGGAGATCTCTCGG	7169	Db	8115	ATGGCCCTTTACATGTGCTCTCCACCTCCCTCAGCGCTGATGGGCTCTTCATACGGA	8174
Qy	2277	ProLysLysArgPheProAlaLeuProAlaThrAlaArgProAspTyrAsnProPro	2296	Qy	2635	PheGlnTyrSerProAlaGlnArgValGluPheLeuLeuLysAlaTrpAlaGluLysLys	2654
Db	7170	AGGTCCAGGAATTTCCCTCGAGCGATGCCATATGAGGACGCCCGATTACACCTTCCA	7229	Db	8175	TTCCAATCTCTCTCGACAGCGGGTCCGAGTTCTCTGTAATGCTTGAAGCGAAGAA	8234
Qy	2297	LeuValGluSerTyrLysArgProAspTyrGlnProAlaThrValAlaGlyCysAlaLeu	2316	Qy	2655	AspProMetGlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluArgAsp	2674
Db	7230	CTGTAGAGTCTGGAGGACCGGACTACGCTCCAGTGGGTACACGGGTGTCATG	7289	Db	8235	TGCGCTATGGGCTTCGCATATGACCCGCTGTTTTGACTCAACGGTCTCTGAGATGAC	8294
Qy	2317	ProProProArgLysThrProThrProProProArgArgArgThrValGlyLeuSer	2336	Qy	2675	IleArgThrGluGluSerIleTyrArgAlaCysSerLeuProGluGluAlaHisThrAla	2694
Db	7290	CCGCTTGCAGAGGCCCTCCCATACCACTCCAGCGGAGGAGGAGCGTGTCTCTGCA	7349	Db	8295	ATCCGTTGTAGAGGTCAATCTACCAATGTTGTGACTTGGCCCCCGAGCAGCAGGCC	8354
Qy	2337	GluAspSerIleGlyAspAlaLeuGlnLeuAlaIleLysSerPheGlyGlnProPro	2356	Qy	2695	IleHisSerLeuThrGluArgLeuTyrValGlyProMetPheAsnSerLysGlyGln	2714
Db	7350	GAATCTACCGTGTCTTCTGCTCTGGCGAGCTCGCCCAAGAGACCTTCGGCAGCTCCGAA	7409	Db	8355	ATAAGGTCTCACAGCGGCTTTACATCGGGGGCCCCCTGACTAATCTTAAAGGGCAG	8414
Qy	2357	ProSerGly---AspSerGlyLeuSerThrGlyAlaGlyAlaAlaAspSerGlySerGln	2375	Qy	2715	ThrCysGlyTyrArgArgCysArgAlaSerGlyValLeuThrThrSerMetGlyAsnThr	2734
Db	7410	TCGTCCGGCGTCCAGCGGACCGGACCGCAACGGCC-----TCTCCTGACGAG	7454	Db	8415	AACTGCGGCTATCGCGGTGCGCGGAGCGGTGTACTGACGACCACTGCGGTATATACC	8474
Qy	2376	ThrProProAspGluLeuAlaLeuSerGluThrGlySerIleSerSerMetProProLeu	2395	Qy	2735	IleThrCysTyrValLysValLeuAlaAlaCysLysAlaGlyIleIleAlaProThr	2754
Db	7455	CCCTCCGACGAGCGGCGGATCCGACGCTTGTGCTGCTTGTGCTTGTGCTTGTGCTTGT	7514	Db	8475	CTCAGATTTACTTGAAGCGCGCTGCGGCTGTGAGCTGCGAGCTCCAGGACTCCAGGCTCAGC	8534
Qy				Qy	2755	MetLeuValCysGlyAspLeuValIleSerGluSerGluGlnGlyThrGluGluAsp	2774

```
Db 4230 GGCAGGTTCTTCCGACCGTGTGCTCTGGGGGGCGCTTATCAGCATCATATATGTGAT 4289
Qy 1321 GluCysHisAlaValAaspSerThrThrIleLeuGlyIleGlyThrValLeuAaspGlnAla 1340
Db 4290 GAGTGCACCTCACTGACTGCACACATATCTGGGATCGGCAGTCTTGGACCAAGCG 4349
Qy 1341 GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProGlySerValThr 1360
Db 4350 GAGACGGCTGGAGCGGACTCGTCTGCTCGCCACCGCTACGCTCCGGGATCGGTCAAC 4409
Qy 1361 ThrProHisProAsnIleGluGluValAlaLeuGlyGlnGluGlyGluIleProPheThr 1380
Db 4410 GTGCCACATCAACATCGAGGAGGTGGCTCTGTCCAGCAGCTCGAGAAATCCCTTTAT 4469
Qy 1381 GlyArgAlaIleProLeuSerThrIleLeuGlyGlyArgHisLeuIlePheCysHisSer 1400
Db 4470 GGCAGACCATCCCATCGAGACCATCAAGGGGGGGAGGACCTCATTTTCTGCCATTC 4529
Qy 1401 LysLysLysCysAaspGluLeuAlaAlaLeuArgGlyMetGlyLeuAenSerValAla 1420
Db 4530 AAGAGAAATGTGATGAGCTCGCGCGAAGCTGTCCGGCTTCGACTCAATGCTGTAGCA 4589
Qy 1421 TyrTyArgGlyLeuAaspValSerValIleProThrGlnGlyAaspValValValAla 1440
Db 4590 TATTACCGGGGCTTGATGATCGTATCGTATACCACTAGCGGAGACGTCATTGTGTAGCA 4649
Qy 1441 ThrAaspAlaLeuMetThrGlyTyThrGlyAaspPheAaspSerValIleAaspCysAenVal 1460
Db 4650 ACGGAGCTCTAATGACGGGCTTTACCGGCGATTTTCGACTCAGTATCGACTCAATACA 4709
Qy 1461 AlaValThrGlnValValAaspPheSerLeuAaspProThrPheThrIleThrGlnIle 1480
Db 4710 TGTGTACCCAGACAGCTTCAGCTTGACCTGGACCCGACCTTCAACATTGAGACGACGACC 4769
Qy 1481 ValProGlnAaspAlaValSerArgSerGlnArgArgGlyArgThrGlyArgGlyArgLeu 1500
Db 4770 GTGCCACAAGACGCGGTGTCACTCGACGCGGAGGAGGAGTGTGTGGGGGAGGATG 4829
Qy 1501 GlyIleTyArgTyValSerThrGlyGluArgAlaSerGlyMetPheAaspSerValVal 1520
Db 4830 GGCATTACAGGTTTGTGATCCAGAGAGAACGCGCCCTCGGGCATGTTCGATTCTCGGTT 4889
Qy 1521 LeuCysGluCysTyAaspAlaGlyAlaAlaTyTyGluLeuThrProSerGluThrThr 1540
Db 4890 CTGTGCGAGTGTATGACGCGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4949
Qy 1541 ValArgLeuArgAlaTyPheAenThrProGlyLeuProValCysGlnAaspHisLeuGlu 1560
Db 4950 GTTAGGTTGGGGCTTACCTTAACACACACAGGCTTGCCTGTCCAGGAGACCATCTGGAG 5009
Qy 1561 PheTrpGluAlaValPheThrGlyLeuThrHisIleAaspAlaHisPheLeuSerGlnThr 1580
Db 5010 TTCTGGGAGAGCTCTTTACAGGCTTACCCATAGACGCGCCATTTCTTGTCCAGACT 5069
Qy 1581 LysGlnSerGlyGluAenPheAlaTyLeuThrAlaTyGlnAlaThrValCysAlaArg 1600
Db 5070 AAGCAGGAGGACAACTTCCCTACCTGTTAGCATACAGGCTACGGTGTGGCCAGG 5129
Qy 1601 AlaLysAlaProProSerTrpAaspValMetTrpLysCysLeuThrArgLeuLysPro 1620
Db 5130 GCTCAGGCTCCACTCCATCGTGGACCAATGTGGAGTGCTCTCATACGGCTAAAGCT 5189
Qy 1621 ThrLeuValGlyProThrProLeuLeuTyArgLeuGlySerValThrAenGluValThr 1640
Db 5190 ACGCTGCAGGGGCAACGCCCTCTGTATAGGCTGGAGCGCTTCAAAACGAGGTTACT 5249
Qy 1641 LeuThrHisProValThrLysTyIleAlaThrCysMetGlnAlaAaspLeuGluValMet 1660
Db 5250 ACCACACACCCCAATACCAATACATCGGCATGTCGCTGACCTGAGGTCGTC 5309
Qy 1661 ThrSerThrTrpValLeuAlaGlyValLeuAlaAlaValAlaAlaTyCysLeuAla 1680
```

Copied from 09540843 on 05/19/2004

Copied from 09540843 on 05/19/2004

7935 AGCAAGCCGCTTAACACATCCGCTCGTGTGGAAGGACTTGTCTGGAAGACACTGAGACA 7994
 2575 ProileProThrThrileMetAlaLysAenGluValPheCysValaAspProThrLysGly 2594
 7995 CCAATTGACACCACTCATGGCAAAATGAGGTTTCTGCGTCCAAACAGAGAAGGGG 8054
 2595 GlyLysLysAlaAlaArgLeuIleValTyrProAspLeuGlyValArgValCysGluLys 2614
 8055 GGCCGCAAGCCAGCTCGCCCTTATCGTATCCAGATTGGGGGTCTGTTGTGCGAGAAA 8114
 2615 MetAlaLeuTyrAspIleThrGlnLysLeuProGlnAlaValMetGlyAlaSerTyrGly 2634
 8115 ATGGCCCTTTACGATGTGGTCTCCACCCCTCCCTCAGCCGCTGATGGGCTCTTCATACGGA 8174
 2635 PheGlnTyrSerProAlaGlnArgValGluPheLeuLeuLysAlaIleAlaGluLysLys 2654
 8175 TTCATATCTCTCTTGACAGCGGGTTCGATGTTCTGTGTAATGCTGGAAGCGAAGAAA 8234
 2655 AspProMetGlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluArgAsp 2674
 8235 TGCCCTATGGCTTCGATATGACACCCGCTGTTTGGACTCAACGGTCACTGAGAAATGAC 8294
 2675 IleArgThrGluGluSerIleTyrArgAlaCysSerLeuProGluGluAlaHisThrAla 2694
 8295 ATCCGTGTTGAGGAGTCAATCTACCAATGTTGTGACTTGGCCCGCCGGAAGCCAGACAGCC 8354
 2695 IleHisSerLeuThrGluArgLeuTyrValGlyProMetPheAsnSerLysGlyGln 2714
 8355 ATAAGGTGCTGTCAGAGCGGCTTACATCGGGGCCCCCTGACTTAATCTTAAGGGCAG 8414
 2715 ThrCysGlyTyrArgArgCysArgAlaSerGlyValLeuThrThrSerMetGlyAsnThr 2734
 8415 AACTGCGGCTATCGCGGTGCGCGCGAGGGTGTACTGACGACCGAGTCCGGTATATACC 8474
 2735 IleThrCysTyrValLysAlaLeuAlaCysLysAlaAlaGlyIleLeuAlaProThr 2754
 8475 CTCACATGTTACTTGAAGGCGCTGCGCGCTGTGAGCTCGAAGCTCCAGGACTGCACG 8534
 2755 MetLeuValCysGlyAspAspLeuValIleSerGluSerGlnGlyThrGluGluAsp 2774
 8535 ATGCTGATGCGGAGACGACCTTGTCTGTATCTGTGAAGCGCGGGAGCCCAAGAGGAC 8594
 2775 GluArgAlaLeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAsp 2794
 8595 GAGCGAGCCTACGGGCTTCACGAGGCTATGACTAGATATCTGCCCCCTCGGGGAC 8654
 2795 ProProArgProGluTyrAspLeuLeuLeuThrSerCysSerSerAsnValSerVal 2814
 8655 CCGCCCAACCCAGAAATACGACTTGGAGTTGATACATCATGCTCTCTCAATGTGTCAATC 8714
 2815 AlaLeuGlyProGlnGlyArgArgTyrTyrLeuThrArgAspProThrThrProIle 2834
 8715 GCGCAGCATGATCGGCAAAAGGGTGTACTATCTACCCGCTGACCCACCCACCCCTT 8774
 2835 AlaArgAlaAlaTrpGluThrValArgHisSerProValAsnSerTrpLeuGlyAsnIle 2854
 8775 GCGCGGGCTGGTGGGAGACAGCTAGACACACTCCAGTCAATCTCTGGCTAGGCAATC 8834
 2855 IleGlnTyrAlaProThrIleTrpAlaArgMetValLeuMetThrHisPheSerIle 2874
 8835 ATCATGTATGGCCCACTTGTGGCAAGGATGATCTGTGATGACTCATCTCTCTCCATC 8894
 2875 LeuMetAlaGlnAspThrLeuAspGlnAsnLeuAsnPheGluMetTyrGlyAlaValTyr 2894
 8895 CTTCTAGCTAGGAACAACTTGAAGAGCCCTAGATTGTCTAGATCTACGGGGCTGTATC 8954
 2895 SerValSerProLeuAspLeuProAlaIleIleGluArgLeuHisGlyLeuAspAlaPhe 2914
 8955 TCCATTGAGCACTTGTACCTTACTTACATCTCAAGCTCCATGGCTTAGCGCATTT 9014
 2915 SerLeuHisThrTyrThrProHisGluLeuThrArgValAlaSerAlaLeuArgLysLeu 2934
 9015 TCACTCCATAGTTACTCTCCAGGTTGAGATCAATAGGTTGCTTCTATGCTCTCAGGAACCTT 9074

2935 GlyAlaProProLeuArgAlaTrpLysSerArgAlaAlaArgAlaValArgAlaSerLeuIle 2954
 9075 GGGGTACCGCCCTTCGAGTCTGGAGACATCGGGCAGAGGTGTCGGCTAGGCTACTG 9134
 2955 SerArgGlyGlyArgAlaAlaValCysGlyArgTyrLeuPheAsnTrpAlaValLysThr 2974
 9135 TCCAGGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTCACTGGGCGAGTAAGGACC 9194
 2975 LysLeuLysLeuThrProLeuProGluAlaArgLeuLeuAspLeuSerSerTrpPheThr 2994
 9195 AAGCTCAAACTCCTCAATCCCGCTGCTGCCAGTTGGATTTATCCAGCTGGTTCGT 9254
 2995 ValGlyAlaGlyGlyAspIleTyrHisSerValSerArgAlaArgProArgLeuLeu 3014
 9255 GCTGGTTACAGCGGGGAGACATATATACAGCTGTCTCGTCCGCGACCCCGCTGGTTC 9314
 3015 LeuPheGlyLeuLeuLeuLeuPheValGlyValGlyLeuPheLeuLeuProAlaArg 3033
 9315 ATGTGGTGTCTACTTCTGTAGGGGTAGGCATCTATCTACTCTCCCAACCGA 9371

RESULT 12

ABK91428
ID ABK91428 standard; DNA; 9605 BP.

XX
AC ABK91428;

XX 15-NOV-2002 (first entry)

XX Hepatitis C virus Con 1 isolate DNA mutant 5.

XX HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;

XX hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;

XX internal ribosome entry site; IRES; NS5A; HCV replication; mutant.

OS Hepatitis C virus.

OS Synthetic.

XX Key Location/Qualifiers

XX CDS 342..9374

XX /tag= a

XX /product= "HCV polyprotein"

XX /note= "The polyprotein consists of the Core, E1,

XX E2, E7, NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"

XX mutation replace (6463,A)

XX /tag= b

XX WO200259321-A2.

XX 01-AUG-2002.

XX 16-JAN-2002; 2002WO-EP00526.

XX 23-JAN-2001; 2001US-263479P.

XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.

XX De Francesco R, Migliaccio G, Paonessa G;

XX WPI; 2002-599793/64.

XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV

XX NS5 encoding region, or encephalomyocarditis virus (EMCV) internal

XX ribosome entry site (IRES) region, useful in studying HCV replication

XX and expression

XX Claim 9; Page -; 69pp; English.

XX The invention relates to nucleic acid molecules comprising altered HCV

XX NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)

XX internal ribosome entry site (IRES) region coding for one or more NS3,

XX NS5A, or EMCV IRES mutations, respectively. The location of the

XX mutations are detailed in the specification. Also included are

1841 AlaAlaValGlySerIleGlyLeuGlyLysValLeuValAspIleLeuAlaGlyTyrGly 1860
Db |||||
5850 GCGCTGTTGGCAGCATAGGCTTGGAGAGTCTTGTGATATTTTGGCAGGTTATGGA 5909
QY |||||
1861 AlaGlyIleSerGlyAlaLeuValAlaPheLysIleMetSerGlyLysProSerMet 1880
Db |||||
5910 GCAGGGGTGGCAGCGCGCTGCTGGCTTTTAAAGTTCATGAGCGGAGATGCCCTCCACC 5969
QY |||||
1881 GluAspValValAsnLeuLeuProGlyIleLeuSerProGlyValAlaLeuValGlyVal 1900
Db |||||
5970 GAGACTGTTTAACTACT 6029
QY |||||
1901 IleCysAlaAlaIleLeuArgHisValGlyProGlyGlyAlaValGlnTrpMet 1920
Db |||||
6030 GTGTGCGCAGCATCTGCTGCGCAGCTGGGCGCCAGGCGAGGCGTGTGCGAGTGGATG 6089
QY |||||
1921 AsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaProThrHisTyrValThr 1940
Db |||||
6090 AACCGGCTGATAGCTTGGCTTCCGCGGGTAAACACGCTCTCCGCCACGCACTATGTGCT 6149
QY |||||
1941 GluSerAspAlaSerGlnArgValThrGlnLeuLeuGlySerLeuThrIleThrSerLeu 1960
Db |||||
6150 GAGAGCAGCTGCAGCAGCTGCTACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 6209
QY |||||
1961 LeuArgArgLeuHisAsnTrpIleThrGluAspCysProIleProCysGlyGlySerTrp 1980
Db |||||
6210 CTGAGAGGCTTCACAGTGGATCAACAGGAGCTGCTCCAGCGCATGCTCGGCTCGTG 6269
QY |||||
1981 LeuArgAspValTrpAspTrpValCysThrIleLeuThrAspPheLysAsnTrpLeuThr 2000
Db |||||
6270 CTAAGAGATGTTTGGGATGGATATGACGGTGTGACTGATTTCAAGACCTGGCTCCAG 6329
QY |||||
2001 SerLeuLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTyrLys 2020
Db |||||
6330 TCCAAAGCTCTGCGCGATGCGGGAGTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 6389
QY |||||
2021 GlyValTrpAlaGlyThrGlyIleMetThrThrArgCysProCysGlyAlaAsnIleSer 2040
Db |||||
6390 GGAGTCTGGCGGCGCAGCGCATCATGCAACACCTGCCCATGTGGAGCAGCATCACC 6449
QY |||||
2041 GlyAsnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTrp 2060
Db |||||
6450 GGACATGTGAAACAGGTTTCCATGAGGATCGTGGGCGCTAGGACCTGTAGTAACAGTGG 6509
QY |||||
2061 GlnGlyThrPheProIleAsnCysTyrThrGluGlyGlnCysValProLysProAlaPro 2080
Db |||||
6510 CATGGAACTTCCCATTAACGGGTACACACGCGGCCCTTGACGCGCTCCCGCGGCCA 6569
QY |||||
2081 AsnPheLysValAlaIleTrpArgValAlaAlaSerGluTyrAlaGluValThrGlnHis 2100
Db |||||
6570 AATTATTTAGGCGCTGTGGCGGTGGCTGTGAGGAGTACGTGGAGGTACGCGGGTG 6629
QY |||||
2101 GlySerTyrHisTyrIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeu 2120
Db |||||
6630 GGGGATTTCCACTACTGACGCGGCATGACCATGACCAACGTAAGTGGCCGTGTGAGT 6689
QY |||||
2121 ProSerProGluPhePheSerTrpValAspGlyValGlnIleHisArgPheAlaProThr 2140
Db |||||
6690 CCGGCCCGGAAATTTCTCAGAAAGTGGATGGGGTGGCTGTGACAGGTACGCTCCAGCG 6749
QY |||||
2141 ProLysProPhePheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal 2160
Db |||||
6750 TGCMAACCCCTCTACGGAGAGGTCATATTCCTGCTGGGCTCTCATCAATACCTGGTT 6809
QY |||||
2161 GlySerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThr 2180
Db |||||
6810 GGGTACAGCTCCATGCGAGCCCGAACCGGAGTACGTGCTCACTTCCATGCTCACC 6869
QY |||||
2181 AspProSerHisIleThrAlaGluThrAlaAlaArgArgLeuAlaArgGlySerProPro 2200
Db |||||
6870 GACCCCTCCACATTAACGGGAGAGCGCTTAAGCTAGGCTGGCCAGGGGATCTCCCCC 6929
QY |||||
2201 SerGluAlaSerSerAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThr 2220

6930 TCTCGGCGCAGCTCATCAGTAGCAGCTGTCTGGCCCTTCTTGAAGCAACATGCACT 6989
QY |||||
2221 ThrHisGlyLysAlaTyrAspValAspMetValAspAlaAsnLeuPhe 2236
Db |||||
6990 ACCCGTTCATGACTCCCGGAGCCTGACCTCATCGAGGCAACCTCTCTGTGGCGGAGGAG 7049
QY |||||
2237 MetGlyGlyAspValThrArgIleGluSerGlySerLysValValValLeuAspSerLeu 2256
Db |||||
7050 ATGGGCGGGAACATCACCCTGCTGAGTCAAAATAGTAGTAAATTTGGACTCTTTC 7109
QY |||||
2257 AspProMetValGluGluArgSerAspLeuProSerIleProSerGluTyrMetLeu 2276
Db |||||
7110 GAGCGCTCCAAGCGAGGAGGATGAGAGGAAGTATCCGTTCCGCGGAGATCTCTGCGG 7169
QY |||||
2277 ProLysLysArgPheProAlaLeuProAlaTyrAlaArgProAspTyrAsnProPro 2296
Db |||||
7170 AGGTCAGGAAATTCCTCGAGCGATGCCATATGGGACGCGCGGATTAACAACCTCCA 7229
QY |||||
2297 LeuValGluSerTrpLysArgProAspTyrGlnProAlaThrValAlaGlyCysAlaLeu 2316
Db |||||
7230 CTGTTAGAGTCTTGGAGACCGGACTACGTCCTCTCCAGTGTACACGGGTGTCCATTG 7289
QY |||||
2317 ProProProArgLysThrProThrProProArgArgArgThrValGlyLeuSer 2336
Db |||||
7290 CCGCTGCCAAGGCCCTCCGATACCATCTCCAGGAGGAGGAGCGGTGTCTCTGTCA 7349
QY |||||
2337 GluAspSerIleGlyAspAlaLeuGlnLeuAlaIleLysSerPheGlyGlnProPro 2356
Db |||||
7350 GAATCTACGCTCTTCTGCTTGGCGAGCTCGCACAAAGACCTTCGCGAGCTCCGAA 7409
QY |||||
2357 ProSerGly--AspSerGlyLeuSerThrGlyAlaGlyAlaAspSerGlySerGln 2375
Db |||||
7410 TCGTCGCGCTCGACGCGGCACGCGCC-----TCTCTGACCAG 7454
QY |||||
2376 ThrProProAspGluLeuAlaLeuSerGluThrGlySerIleSerSerMetProProLeu 2395
Db |||||
7455 CCTCCGAGCAGCGCGGATCCGAGCTTGGTGTGACTCTCTCCATGCCCCCTT 7514
QY |||||
2396 GluGlyGluLeuGlyAspProAspLeuGluProGluGlnValGluProGlnProPro 2415
Db |||||
7515 GAGGGGAGCGCGGGGATCCCGATCTC----- 7541
QY |||||
2416 GlnGlyGlyValAlaAlaProGlySerAspSerGlySerTrpSerThrCysSerGluGlu 2435
Db |||||
7542 -----AGCGAGGGTCTTGGTCTTACCGTAGCGAGGAG 7574
QY |||||
2436 ---AspAspSerValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuIleThr 2454
Db |||||
7575 GCTAGTGAGGAGCTGCTGCTGCTCGATGCTCTACATGACAGGCGCCCTGATCAG 7634
QY |||||
2455 ProCysSerProGluGluGlyLysLeuProIleAsnProLeuSerAsnSerLeuLeuArg 2474
Db |||||
7635 CCATCGCTGCGGAGAAACCAAGCTGCCCATCAATGCACTGAGCAACTCTTTGTCTCGT 7694
QY |||||
2475 TyrHisAsnLysValTyrCysThrThrThrLysSerAlaSerLeuArgAlaLysVal 2494
Db |||||
7695 CACCAACTTGGTCTATGCTCAACATCTCCGACGCAAGCTCGCGCAGAGGAGGTC 7754
QY |||||
2495 ThrPheAspArgMetGlnValLeuAspSerTyrTyrAspSerValLeuLysAspIleLys 2514
Db |||||
7755 ACCTTTGACAGACTCAGGTCTGCGACCACTACCGGAGCTGCTCAAGGAGATGAAG 7814
QY |||||
2515 LeuAlaIleSerLysValThrAlaArgLeuLeuMetGluGluAlaCysGlnLeuThr 2534
Db |||||
7815 GCGAGGCGCTCCACAGCTTAAAGCTTAACTTCTATCCGTGGAGGAAGCTGTAGCTGAGC 7874
QY |||||
2535 ProProHisSerAlaArgSerLysTyrGlyPheGlyAlaLysGluValArgSerLeuSer 2554
Db |||||
7875 CCCCCACATTCGCGCAGATCTAAATTTGGTATGGGCAAGAGCTCCGGAACCTATCC 7934
QY |||||
2555 GlyArgAlaValAsnHisIleLysSerValTrpLysAspLeuLeuGluAspSerGluThr 2574
Db |||||

Db 3630 ATCACCAAAATGTACCAATGTGGACAGGACCTCGTGGCTGGCAAGCGCCCGCGG 3689
 Qy 1121 ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuThrArgAsn 1140
 Db 3690 GCGCGTTCCTTGTACACATGACCTCGGCGAGCTCGGACCTTTACTTGGTCAGAGCAT 3749
 Qy 1141 AlaAspValIleProAlaArgArgArgGlyAspArgGlyAlaLeuLeuSerProArg 1160
 Db 3750 GCGCATGTCTATCGGTCGCGCGGCGGCGAGCAGCAGGGGAGCTACTCTCCCGCAG 3809
 Qy 1161 ProLeuSerThrLeuLysGlySerSerGlyProValLeuCysProArgGlyHisAla 1180
 Db 3810 CCGCTCTCTTACTTGAAGGCTCTCGGCGGTCCACTGTCTGCGCCCTCGGCGCACGCT 3869
 Qy 1181 ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle 1200
 Db 3870 GTGGGATCTTTCGGGCTGCGGTGCGCCGAGGGGTTCGAGGCGGTGAGCTTTGTA 3929
 Qy 1201 ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro 1220
 Db 3930 CCGCTCGAGTCTATGGAAACCACTATCGGCTCCCGGTCTTCACGGACAACTCGTCCCT 3989
 Qy 1221 ProAlaValProGluThrThrGluValGlyThrLeuHisAlaProThrGlySerGlyLys 1240
 Db 3990 CCGCGCTACCGCAGACATTCAGGTGGCCCTATCACCGCCCTACTGTGTAGCGCAAG 4049
 Qy 1241 SerThrLysValProValAlaThrAlaAlaGluGlyThrLysValLeuValLeuAsnPro 1260
 Db 4050 AGCACTAAGGTGCGGCTGCGTATGACAGCCAGGCTTAAGGTGCTTGTCTGAACCCG 4109
 Qy 1261 SerValAlaAlaThrLeuGlyPheGlyAlaThrLysSerLysAlaHisGlyIleAsnPro 1280
 Db 4110 TCGCTCGCCGCCACCTAGTTTCGGGGCGTATATGCTAAGGCACATGATCGACCT 4169
 Qy 1281 AsnIleArgThrGlyValAlaGthrValThrThrGlyAlaProIleThrSerThrTyr 1300
 Db 4170 AACATCAGAACCGGGTAAGGACCATCACACGGGTGCCCGCCCATCAGTACTCCACCTAT 4229
 Qy 1301 GlyLysPheLeuAlaAspGlyGlyCysAlaGlyValAlaThrAspIleIleCysAsp 1320
 Db 4230 GGCAAGTTTCTTGGCGAGCGGTGCTCTGGGGCGCCCTATGACATCAATATGTAT 4289
 Qy 1321 GluCysHisAlaAspSerThrIleLeuGlyIleGlyThrValLeuAspGluAla 1340
 Db 4290 GAGTGCACCACTCACTGACACCATCTCTGGGCATCGGCACAGTCTCGACCAAGC 4349
 Qy 1341 GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProGlySerValThr 1360
 Db 4350 GAGACGGCTGGAGCGGACCTGCTGCTGCTGCCACCGCTACCGCTCCGGGATCGTCA 4409
 Qy 1361 ThrProHisProAsnIleGluGluValAlaLeuGluGluGlyIleProPheTyr 1380
 Db 4410 GTGCCACATCCAAACATCGAGGAGGTGGCTCTGTCCAGCACTGGAGAAATCCCTTTAT 4469
 Qy 1381 GlyArgAlaIleProLeuSerTyrIleLysGlyGlyArgHisLeuIlePheCysHisSer 1400
 Db 4470 GGCAAGCCATCCCATCGACACCATCAAGGGGGGGAGGCACCTCACTTTCTGCTATCC 4529
 Qy 1401 LysLysLysCysAspGluLeuAlaAlaAlaLeuArgGlyMetGlyLeuAsnSerValAla 1420
 Db 4530 AAGAAGAAATGTGATGAGTGGCGCGGAGGTGTCCGCGCTCGGACCTCAATGTGTAGCA 4589
 Qy 1421 TyrThrArgGlyLeuAspValSerValIleProThrGluGlyAspValValValAla 1440
 Db 4590 TATTACCGGGGCTTGTATGATTCGTCATACCACTAGCGGAGACGTCATTGTGCTAGCA 4649
 Qy 1441 ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnVal 1460
 Db 4650 ACGGACGCTCTAATGACGGGCTTTACCGCGATTTCCGACTCAGTGTGACTGCAATACA 4709
 Qy 1461 AlaValThrGluValAlaAspPheSerLeuAspProThrPheThrIleThrThrGluIle 1480
 Db 4710 TGTGTACCCAGACAGTCTGACTTCAGCGCTGACCCGACCTTCACCATTTGAGACGACGACC 4769

Qy 1481 ValProGlnAspAlaValSerArgSerGlnArgArgGlyArgThrGlyArgGlyArgLeu 1500
 Db 4770 GTGCCCAAGAGCGCGGTGTCAAGCTCGCAGCGCGGAGGAGGACTGGTAGGGGAGGATG 4829
 Qy 1501 GlyIleTyrArgTyrValSerThrGlyGluArgAlaSerGlyMetPheAspSerValVal 1520
 Db 4830 GGCATTTACAGGTTTGTGACTCCAGGAGAACCGCCCTCGGGCATGTTCGATTTCTCGGTT 4889
 Qy 1521 LeuCysGluCysTyrAspAlaGlyAlaAlaTyrTyrGluLeuThrProSerGluThrThr 1540
 Db 4890 CTGTGCGAGTGTATGACCGCGGCTGTCTTGGTAGAGCTCACGCCCGCCGAGACCTCA 4949
 Qy 1541 ValArgLeuArgAlaTyrPheAsnThrProGlyLeuProValCysGlnAspHisLeuGlu 1560
 Db 4950 GTTAGGTTGCGGCTTACCTAAACACACCCAGGCTTCCCGCTCTGCCAGGACCATCTGGAG 5009
 Qy 1561 PheTrpGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThr 1580
 Db 5010 TTCTGGGAGAGCGTCTTTACAGGCTCACCCACATAGACGCCCATTTCTTGTCCAGACT 5069
 Qy 1581 LysGlnSerGlyGluAsnPheAlaTyrLeuThrAlaTyrGlnAlaThrValCysAlaArg 1600
 Db 5070 AAGCAGGAGGAGAACCTTCCCTACCTGCTGTGTAGCATACAGGCTACGGTGTGGCCAGG 5129
 Qy 1601 AlaLysAlaProProProSerTrpAspValMetTrpLysCysLeuThrArgLeuLysPro 1620
 Db 5130 GCTCAGGCTCCACCTCCATCGTGGGACCAATGTGGAAGTGTCTCATACGGCTAAAGCCT 5189
 Qy 1621 ThrLeuValGlyProThrProLeuLeuTyrArgLeuGlySerValThrAsnGluValThr 1640
 Db 5190 ACGCTCAGCGGCAACGCCCTTCTGTGTATAGGCTGGGAGCGGTTTCAAAACGAGGTTACT 5249
 Qy 1641 LeuThrHisProValThrLysTyrIleAlaThrCysMetGlnAlaAspLeuGluValMet 1660
 Db 5250 ACCACACACCCCATACCAATACATCATGGCATGTCATCGGTGAGCTTGGAGGTGCTC 5309
 Qy 1661 ThrSerThrTrpValLeuAlaGlyGlyValLeuAlaAlaValAlaAlaTyrCysLeuAla 1680
 Db 5310 ACGAGCACCTGGGTGCTGTAGGCGGAGTCTTAGCAGCTCTGGCCCGGATTTCCTGACA 5369
 Qy 1681 ThrGlyCysValCysIleIleGlyArgLeuHisIleAsnGlnArgAlaValAlaAlaPro 1700
 Db 5370 ACAGGACGCGTGTCTATTTGGGCGAGGATCATCTTGTCCGGAAGCGGCCCATCATCC 5429
 Qy 1701 AspLysGluValLeuGluAlaPheAspGluMetGluCysAlaSerArgAlaAla 1720
 Db 5430 GACAGGAGTCTCTTTTACCGGGAGTTTCGATGAGATGGAAGAGTGGCGCTCACACCTCC 5489
 Qy 1721 LeuIleGluGluGlyGlnArgIleAlaGluMetLeuLysSerLysIleGlnGlyLeuLeu 1740
 Db 5490 TACATCGAACAGGAATGAGCTCGCGGAACCAATTCAAACAGAGGCAATCGGGTTGCTG 5549
 Qy 1741 GlnGlnAlaSerLysGlnAlaGlnAspIleGlnProThrValGlnAlaSerTrpProLys 1760
 Db 5550 CAAACAGCCCAAGCAAGCGGAGGTGCTGCTCCCGTGGTGAATCCCAAGTGGCGGACC 5609
 Qy 1761 ValGluGlnPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeuAla 1780
 Db 5610 CTGGAAGCCTTCTGGGGGAGCATATGTGAATTTTCATCAGCGGATACATATTTAGCA 5669
 Qy 1781 GlyLeuSerThrLeuProGlyAsnProAlaValAlaSerMetMetAlaPheSerAlaAla 1800
 Db 5670 GGCTGTGCTCCTGCTGCGCAACCCCGCATAGCATCATGTATGTCATTCACAGCCTCT 5729
 Qy 1801 LeuThrSerProLeuSerThrSerThrThrIleLeuLeuAsnIleLeuGlyTyrLeu 1820
 Db 5730 ATCAGGCGCGCTCACCAACCAATCCCTCTCTCTTTAAACATCTCTGGGGGATGGGTG 5789
 Qy 1821 AlaSerGlnIleAlaProAlaGlyAlaThrGlyPheValValSerGlyLeuValGly 1840
 Db 5790 GCCGCCCAACTGCTCTCCAGCGGTGCTTCTGCTTTCGTAGGCGCGGCGCATCGCTGGA 5849

QY 381 ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr 400
Db 1482 GTTACGGGGAAACCTATGTGACAGGGGGGACGATGGCCAAAACACACCTCGGGATTACG 1541
QY 401 SerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTrp 420
Db 1542 TCCTCTCTTTTACCCCGGGTCATCCAGAAATCCAGCTTGTAAACCAACGCGACGTGG 1601
QY 421 HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer 440
Db 1602 CACATCACAGGAGTGCCTTCACTGCAATGATCTCCCTCAACACTGGTTCCTTGTGCG 1661
QY 441 LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg 460
Db 1662 CTGTCTTACGTGCACAAAGTTCACTCATCTGGATGCCAGAGCGCATGGCCAGCTGCAGC 1721
QY 461 SerIleGluAlaPheArgValGlyTrpGlyAlaLeuGlnTyrGluAspAsnValThrAsn 480
Db 1722 CCCATCGACGCGTTCGTACAGGGGTGGGGGCCCATCACTTAC-----AATGAGTCACAC 1775
QY 481 ProGluAspMetArgProTyrCysTrpHisTyrProProArgGlnCysGlyValValSer 500
Db 1776 AGCTCGACAGAGGCTTATTTGTGGCACTACGCACCCCGCGCGGTATCGTACCC 1835
QY 501 AlaLysThrValCysGlyProValTyrCysPheThrProSerProValValValGlyThr 520
Db 1836 GCGGCGCAGGTGTGTCCAGTGACTGTCTTCAACCCCAAGCCCTGTCTGTGTGGGAGC 1895
QY 521 ThrAspArgLeuGlyAlaProThrTyrThrTyrTrpGlyGluAsnGluThrAspValPheLeu 540
Db 1896 ACCGACCGGTTCGGCGCTCCCTACGTACGTGGGGGAGAAATGAGACGGACGTGTGCTT 1955
QY 541 LeuAsnSerThrArgProProLeuGlySerTrpPheGlyCysThrTrpMetAsnSerSer 560
Db 1956 CTTAACACACGCGCGCCGCAAGCAACTGTTTGGCTGTACATGGATGAATAGCACT 2015
QY 561 GlyTyrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer 580
Db 2016 GGGTTTCAACCAAGACGTGCGGGGGCCCCCGGTGAACATCGGGGGATCGGCAATAAAACC 2075
QY 581 ThrAspLeuLeuCysProThrAspCysPheArgLysHisProAspThrThrTyrLeuLys 600
Db 2076 -----TTGACCTTGCCTCCAGCACTGTTCCGGAAGCACCCCGAGGCCACTTACACCAAG 2129
QY 601 CysGlySerGlyProThrLeuThrProArgCysLeuIleAspTyrProTyrArgLeuTrp 620
Db 2130 TGTGGTTTCGGGGCTTGGTTGACACCCAGATGCTTGGTCCACTTACCATACAGGCTTTGG 2189
QY 621 HisTyrProCysThrValAsnTyrThrIlePheLysIleArgMetTyrValGlyGlyVal 640
Db 2190 CACTACCTTGCACCTGTCACTTTTACATTTCAAGTTAGGATGTACGTGGGGGAGTG 2249
QY 641 GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp 660
Db 2250 GAGCACAGGCTCGAAGCCGCATGCAATGGACTCGAGAGAGCGTGTAACTTGGAGGAC 2309
QY 661 ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGluTrpAlaIleLeuPro 680
Db 2310 AGGGACAGATCAGAGCTTACCGCGCTGCTGCTCTACAAACGAGGTGGCAGGTATTGGCC 2369
QY 681 CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisLeuHisGlnAsnIle 700
Db 2370 TGTTCCTTCAACACCTTACCGGCTGTCTCACTGGTTTGTATCCATCTCCATCAGACGTC 2429
QY 701 ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrLysTyrIleValArgTrp 720
Db 2430 GTGACGCTCAATACCTGTACGCTATAGGTCGCGCTGCTCTCTCTTGTCAATCAATGG 2489
QY 721 GluTrpValIleLeuPheLeuLeuLeuAlaAspAlaArgValCysAlaCysLeuTrp 740
Db 2490 GAGTATGCTGTGTCTTCTTCT 2549
QY 741 MetLeuIleLeuLeuGlyGlnAlaGluAlaAlaLeuLeuLysLeuValIleLeuHisAla 760

Db 2550 ATGATGCTCTGATAGCTCAAGCTGAGCGCGCCTAGAGAAACCTGTGTGTCTCTCAACGCG 2609
QY 761 AlaSerAlaAlaSerCysAsnGlyPheLeuTyrPheValIlePhePheValAlaAlaTrp 780
Db 2610 GCATCCGTGGCGGGCGCATGGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2669
QY 781 TyrIleLysGlyArgValValProLeuAlaThrTyrSerLeuThrGlyLeuTrpSerPhe 800
Db 2670 TACATCAGCGGACGCTGTCTCTCTGGGGCGCATATGCTCTCTACGGGTATGGCGCTA 2729
QY 801 SerLeuLeuLeuLeuAlaLeuProGlnGlnAlaTyrAlaTyrAspAlaSerValHisGly 820
Db 2730 CTCTCTCTCTCTCTGGCGTTTACCACGAGCATACGCGCATGCGACGGGAGATGGCGCA 2789
QY 821 GlnIleGlyAlaAlaLeuLeuValMetIleThrLeuPheThrLeuThrProGlyTyrLys 840
Db 2790 TCGTGGGAGGCGCGGTTCGTAGGTCTGATACCTCTTGACCTTGTACCGCATATAAG 2849
QY 841 ThrLeuLeuSerArgPheLeuTrpTyrLeuCysTyrLeuLeuThrLeuGlyGluAlaMet 860
Db 2850 CTGTTCTCTGATAGGTCAATATGGTGTACAAATATTTATCACAGGGCGGAGGCACAC 2909
QY 861 ValGlnGluTrpAlaProProMetGlnValArgGlyGlyArgAspGlyIleIleTrpAla 880
Db 2910 TTGCAAGTGTGGATCCCCCTCAAGCTTCGGGGGGCGCGATGCGCTCATCTCTCTC 2969
QY 881 ValAlaIlePheTyrProGlyValValPheAspIleThrLysTrpLeuLeuAlaValLeu 900
Db 2970 ACGTGGCGGATCCACCCAGAGCTAATCTTTACCATCACCAAAATCTTGTCTCGCATACT 3029
QY 901 GlyProAlaTyrLeuLysGlyAlaLeuThrArgValProTyrPheValArgAlaHis 920
Db 3030 GGTCCACTCATGCTGCTCCAGCTGTATACCAAGTCCGCTACTTCTGTGGCGGCACAC 3089
QY 921 AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyGlyArgTyrValGlnMet 940
Db 3090 GGGCTCATTCGTGTCATGTCATGCTGGTGGGAAAGTTGCTGGGGTCTATTATGTCCAAATG 3149
QY 941 AlaLeuLeuAlaLeuGlyArgTrpThrGlyThrTyrIleTyrAspHisLeuThrProMet 960
Db 3150 GCTCTCATGAAGTTGGCGCACTGACAGGTACGTAGTATGATGATGATGATGATGATGAT 3209
QY 961 SerAspTrpAlaAlaSerGlyLeuArgAspLeuAlaValAlaGluProIleIlePhe 980
Db 3210 CCGGACTGGGCCACGCGGGCTTACAGACCTTGGTGGCAGTTGAGCCGCTGTCTTC 3269
QY 981 SerProMetGluLysLysValIleValTrpGlyAlaGluThrAlaAlaCysGlyAspIle 1000
Db 3270 TCTGATATGAGACCAAGGTTATCCTGGGGGGCAGACACCGCGCGTGTGGGACATC 3329
QY 1001 LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp 1020
Db 3330 ATCTTGGGCTCTCGCGCTCTCCCGCGAGGGAGGAGATACATCTGGGACCGGAGAC 3389
QY 1021 GlyTyrThrSerLysGlyTrpSerLeuLeuAlaProIleThrAlaTyrAlaGlnGlnThr 1040
Db 3390 AGCCTTGAAGGCGAGGGTGGCGACTCTCTCGCGCTTATACGGCTTACTCCCAACAGACG 3449
QY 1041 ArgGlyLeuLeuGlyThrIleValSerMetThrGlyArgAspLysThrGluGlnAla 1060
Db 3450 CGAGGCTTACTTGGCTGTCATCATCTAGCTCACAGGCGCGGACAGGAACACGAGTCGAG 3509
QY 1061 GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly 1080
Db 3510 GGGGAGGTCAAGTGTCTTCCACCGCAACACAATCTTCTCTGGCGACCTCGCTGCTCAATGG 3569
QY 1081 ValLeuTrpThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro 1100
Db 3570 GTGTGTGAGTGTCTATCATGTGCTCCGCTCAAGACCTTCTGGCGGCCCAAGGGGCCCA 3629
QY 1101 ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTrpProSerProGly 1120

XX Hepatitis C virus.
OS Synthetic.
XX Key Location/Qualifiers
FH CDS 342..9374
FT /*tag= a
FT /product= "HCV polyprotein"
FT /note= "The polyprotein consists of the Core, E1,
FT E2, P7, NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
FT mutation replace (6934,T)
FT /*tag= b
XX WO200259321-A2.
XX
XX
XX PD 01-AUG-2002.
XX
XX PF 16-JAN-2002; 2002WO-EP00526.
XX
XX PR 23-JAN-2001; 2001US-263479P.
XX
XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
XX De Francesco R, Migliaccio G, Paonessa G;
XX WPI; 2002-599793/64.
XX
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
XX NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
XX ribosome entry site (IRES) region, useful in studying HCV replication
XX and expression
XX
XX Claim 9; Page -: 69pp; English.
XX
XX The invention relates to nucleic acid molecules comprising altered HCV
XX NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
XX internal ribosome entry site (IRES) region coding for one or more NS3,
XX NS5A, or EMCV IRES mutations, respectively. The location of the
XX mutations are detailed in the specification. Also included are
XX (1) an expression vector comprising a nucleotide sequence coding for
XX the altered nucleic acids, which is transcriptionally coupled to an
XX exogenous promoter; (2) a recombinant cell human hepatoma cell comprising
XX the altered nucleic acids; (3) a recombinant cell produced by introducing
XX into a human hepatoma cell the altered nucleic acids; (4) producing an
XX HCV (hepatitis C virus) replicon enhanced cell or which containing a
XX functional HCV replicon; (5) an HCV replicon enhanced cells made in the
XX method; and (6) measuring the ability of a compound to affect HCV
XX activity. The HCV replicons and HCV replicon enhanced cells are useful in
XX studying HCV replication and expression, and HCV and host cell
XX interactions, producing HCV RNA and proteins, and providing a system
XX for measuring the ability of a compound to modulate one or more HCV
XX activities e.g. to discover drugs which may treat HCV mediated
XX diseases such as liver failure, cirrhosis and hepatocellular carcinoma.
XX The present sequence is an HCV replicon Con 1 mutant of the invention.
XX Note: The present sequence is not shown in the specification but
XX was created by the indexer using the HCV sequence appearing as
XX ABK91411 and the information in Claim 9.
XX
XX SQ Sequence 9605 BP; 1910 A; 2884 C; 2733 G; 2078 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 0 Length: 9605
XX Score: 12034.50 Matches: 2183
XX Percent Similarity: 83.42% Conservative: 352
XX Best Local Similarity: 71.83% Mismatches: 469
XX Query Watch: 74.66% Indels: 35
XX DB: 24 Gaps: 7
XX
XX US-09-980-559-2 (1-3033) x ABK91431 (1-9605)
XX
XX 1 MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGln 20
XX
XX 342 ATGAGCAGCAATCTAAACCTCAAGAAACCAACCAAGTAACCAACCGCGCCACAG 401

21 AspValLysPheProGlyGlyGlnIleValGlyGlyValTyrIleuLeuProArgArg 40
402 GAGCTCAAGTTCGCGGGCGTGTGATCGTGGAGTTTACCTGTTCGCGCGAG 461
41 GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly 60
462 GCGCCCGAGGTGGGTGGTGGCGCGACTAGAAAGACTTCGAGCGGTGCGAAGCTCGTGA 521
61 ArgArgGlnProIleProLysAspArgArgSerThrGlyLysSerTrpGlyLysProGly 80
522 AGCGCAACCTATCCCAAGGCTGCCAGCCCGAGGTAGGGCTGGGCTAGCCCGGG 581
81 TyrProTrpProLeuTyrGlyAsnGluGlyLeuGlyTrpAlaGlyTrpIleuLeuSerPro 100
582 TACCCCTGCGCCCTCTATGGCAATGAGGCTTGGGTGGCGAGGATGGCTCTGTCAACC 641
101 ArgGlySerArgProSerTrpGlyProAsnAspProArgHisArgSerArgAsnValGly 120
642 CGTGGCTCTCGGCTAGTTGGGGCGCCACGAGCCCGCGGTAGGTTCGGCGCAATTTGGGT 701
121 LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProValVal 140
702 AAGGTATCATGATACCTCAGTGGGCTTCCCGCATCTCATGGGTACATTCGCTCGTC 761
141 GlyAlaProLeuGlyGlyValAlaArgAlaLeuAlaHisGlyValArgValLeuGluAsp 160
762 GCGCCCGCTTAGGGGGCGCTGCCAGGCGCTGGCGCATGGGTTCGGGTTCGGAGGAC 821
161 GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla 180
822 GCGCTGAACATGCAACAGGGAATCTGCCGGTGTCTCTTCTATCTCTCTTTGGCT 881
181 LeuLeuSerCysIleThrThrProValSerAlaAlaGluValLysAsnIleSerThrGly 200
882 TTGCTGTCTGTGTGGTACCATCCAGCTTCGCTTATGAGTGGCGCAAGTATCCGAGTG 941
201 TyrMetValThrAsnAspCysThrAsnAspSerIleThrTrpGlnLeuGlnAlaVal 220
942 TACCATGTCAAGAACGACTGCTCAACGCAAGCATTTGTGTATGAGCGCGGACATGATC 1001
221 LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTrpIle 240
1002 ATCATATCCCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1061
241 ProValSerProAsnValAlaValGlnArgProGlyAlaLeuThrGlnGlyLeuArgThr 260
1062 GCGCTCACTCCACGCTCG 1121
261 HisIleAspMetValValMetSerAlaThrLeuLeuCysSerAlaLeuTyrValGlyAspLeu 280
1122 CATGTGATTTGCTGCTGGGGGGGCTGCTCTCTGCTCGCTATGTAGCTGGGAGATCTC 1181
281 CysGlyGlyValMetLeuAlaGlnMetPheIleValSerProGlnHisHisTrpPhe 300
1182 TGGGATCTGTTTCTCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1241
301 ValGlnAspCysAsnCysSerIleThrIleThrGlyHisArgMetAlaTrp 320
1242 GTACAGGACTGCAATTTGCTCAATATATCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1301
321 AspMetMetMetAsnTrpSerProThrAlaThrMetIleLeuAlaTyrAlaMetArgVal 340
1302 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1361
341 ProGluValIleIleAspIleIleSerGlyAlaHisTrpGlyValMetPheGlyLeuAla 360
1362 CCACAAGCTGTGCTGGATATGGTGGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1421
361 TyrPheSerMetGlnGlyAlaTrpAlaLysValValIleLeuLeuLeuAlaGly 380
1422 TACTATTCATGTTGGGGAACCTGGGCTAAGGTTCTGATGTGATGTACTTCTTTGCCGC 1481

2377	ProProAspGluLeuAlaLeuSerGluThrGlySerIleSerSerMetProProLeuGlu	2396	ProProAspGluLeuAlaLeuSerGluThrGlySerIleSerSerMetProProLeuGlu
7482	CCCCCGCAC-----TCCGACGTTGAGTCTCTATTCTTCCATGCCCCCCTGGAG	7529	CCCCCGCAC-----TCCGACGTTGAGTCTCTATTCTTCCATGCCCCCCTGGAG
2397	GlyGluLeuGlyAspProAspLeuGluProGluGlnValGluProGlnProProGln	2416	GlyGluLeuGlyAspProAspLeuGluProGluGlnValGluProGlnProProGln
7530	GGGGAGCTGGGGATCCGGATCTC-----	7553	GGGGAGCTGGGGATCCGGATCTC-----
2417	GlyGlyValAlaAlaProGlySerAspSerGlySerTrpSerThrCysSerGlu	2434	GlyGlyValAlaAlaProGlySerAspSerGlySerTrpSerThrCysSerGlu
7554	-----AGCGACGGGTCTGTCGACGGTCTAGTACTGGGGCC	7589	-----AGCGACGGGTCTGTCGACGGTCTAGTACTGGGGCC
2435	GluAspAspSerValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuThr	2454	GluAspAspSerValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuThr
7590	GAACCGGAAGATGTCTGTGCTGCTCAATGTCTTATTCTTGACAGCGCACTCGTCACC	7649	GAACCGGAAGATGTCTGTGCTGCTCAATGTCTTATTCTTGACAGCGCACTCGTCACC
2455	ProCysSerProGluGluGlyLeuProIleAsnProLeuSerAsnSerLeuLeuArg	2474	ProCysSerProGluGluGlyLeuProIleAsnProLeuSerAsnSerLeuLeuArg
7650	CCGTGGCTGCGAAGACAATACTGCCATCAACGCACTGAGCAACTCGTTGTCTACGC	7709	CCGTGGCTGCGAAGACAATACTGCCATCAACGCACTGAGCAACTCGTTGTCTACGC
2475	TyrHisAsnLysValTyrCysThrThrThrLysSerAlaSerLeuArgAlaLysVal	2494	TyrHisAsnLysValTyrCysThrThrThrLysSerAlaSerLeuArgAlaLysVal
7710	CATCAAAATCGTGATTCCACCACTTCACGCAGTGTTCGCAAGGCGAGAAGAAGTC	7769	CATCAAAATCGTGATTCCACCACTTCACGCAGTGTTCGCAAGGCGAGAAGAAGTC
2495	ThrPheAspArgMetGlnValLeuAspSerTyrTyrAspSerValLeuLysAspLeuLys	2514	ThrPheAspArgMetGlnValLeuAspSerTyrTyrAspSerValLeuLysAspLeuLys
7770	ACATTTGACAGACTCGCAAGTCTTGGACAGCAATTACCAGGACGTGCTCAAGGAGGTCAA	7829	ACATTTGACAGACTCGCAAGTCTTGGACAGCAATTACCAGGACGTGCTCAAGGAGGTCAA
2515	LeuAlaAlaSerLysValThrAlaArgLeuLeuThrMetGluGluAlaCysGlnLeuThr	2534	LeuAlaAlaSerLysValThrAlaArgLeuLeuThrMetGluGluAlaCysGlnLeuThr
7830	GCAGCGCGCTCAAAAGTGAAGCTTAATCTGCTATCCGTAGAGGAAGCTTGACCTTGACG	7889	GCAGCGCGCTCAAAAGTGAAGCTTAATCTGCTATCCGTAGAGGAAGCTTGACCTTGACG
2535	ProProHisSerAlaArgSerLysTyrGlyPheGlyAlaLysGluValArgSerLeuSer	2554	ProProHisSerAlaArgSerLysTyrGlyPheGlyAlaLysGluValArgSerLeuSer
7890	CCCCCACATTCAGCCAAATCCAAAGTTTGGCTATGGGGCAAAAGACGTCTGTTGCCATGCC	7949	CCCCCACATTCAGCCAAATCCAAAGTTTGGCTATGGGGCAAAAGACGTCTGTTGCCATGCC
2555	GlyArgAlaValAsnHisIleLysSerValTrpLysAspLeuLeuGluAspSerGluThr	2574	GlyArgAlaValAsnHisIleLysSerValTrpLysAspLeuLeuGluAspSerGluThr
7950	AGAAAGCGCGTAGCCCCACATCAACTCCGTGTGGAAAGACCTTCTGGAACACAGTGTAA	8009	AGAAAGCGCGTAGCCCCACATCAACTCCGTGTGGAAAGACCTTCTGGAACACAGTGTAA
2575	ProIleProThrThrIleMetAlaLysAsnGluValPheCysValAspProThrLysGly	2594	ProIleProThrThrIleMetAlaLysAsnGluValPheCysValAspProThrLysGly
8010	CCAAATGACACTACCATCATGTCGCAAGAACAGAGGTTTTCTGCTTCAGCTCAGAGAGGG	8069	CCAAATGACACTACCATCATGTCGCAAGAACAGAGGTTTTCTGCTTCAGCTCAGAGAGGG
2595	GlyLysLysAlaAlaArgLeuIleValTyrProAspLeuGlyValArgValCysGluLys	2614	GlyLysLysAlaAlaArgLeuIleValTyrProAspLeuGlyValArgValCysGluLys
8070	GGTCGTAAGCCAGCTCGTCTCATCGTGTTCGCCGACCTGGGCGTGCCTGTGCCAGAA	8129	GGTCGTAAGCCAGCTCGTCTCATCGTGTTCGCCGACCTGGGCGTGCCTGTGCCAGAA
2615	MetAlaLeuTyrAspIleThrGlnLysLeuProGlnAlaValMetGlyAlaSerTyrGly	2634	MetAlaLeuTyrAspIleThrGlnLysLeuProGlnAlaValMetGlyAlaSerTyrGly
8130	ATGGGCCCTGTACGACGTGGTTAGCAAGCTCCCGCTGGCGCTGATGGGAAGCTCTCA	8189	ATGGGCCCTGTACGACGTGGTTAGCAAGCTCCCGCTGGCGCTGATGGGAAGCTCTCA
2635	PheGlnTyrSerProAlaGlnArgValGluPheLeuLeuLysAlaTrpAlaGluLysLys	2654	PheGlnTyrSerProAlaGlnArgValGluPheLeuLeuLysAlaTrpAlaGluLysLys
8190	TTCCAAATCTACACAGGACAGCGGGTTGAAATTCCTCGTCAAGCTGGAACTCCAA	8249	TTCCAAATCTACACAGGACAGCGGGTTGAAATTCCTCGTCAAGCTGGAACTCCAA
2655	AspProMetGlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluArgAsp	2674	AspProMetGlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluArgAsp
8250	ACCCCGATGGGGTTCTCGTATGATACCCCGCTGTTTGACTCCACAGTCACTGAGAGCG	8309	ACCCCGATGGGGTTCTCGTATGATACCCCGCTGTTTGACTCCACAGTCACTGAGAGCG
2675	IleArgThrGluLysSerIleTyrArgAlaCysSerLeuProGluGluAlaHisThrAla	2694	IleArgThrGluLysSerIleTyrArgAlaCysSerLeuProGluGluAlaHisThrAla
8310	ATCCCGTAGGAGGAGCAATTTACCAATGTTGTGACCTGGACCCCAAGCCCGCTGGCC	8369	ATCCCGTAGGAGGAGCAATTTACCAATGTTGTGACCTGGACCCCAAGCCCGCTGGCC
2695	IleHisSerLeuThrGluArgLeuTyrValGlyGlyProMetPheAsnSerLysGlyGln	2714	IleHisSerLeuThrGluArgLeuTyrValGlyGlyProMetPheAsnSerLysGlyGln
8370	ATCAAGTCCCTCACTGAGNGGCTTATGTTGGGGGCCCTTCTTACCAATTCAGAGGGG	8429	ATCAAGTCCCTCACTGAGNGGCTTATGTTGGGGGCCCTTCTTACCAATTCAGAGGGG
2715	ThrCysGlyTyrArgArgCysArgAlaSerGlyValLeuThrThrSerMetGlyAsnThr	2734	ThrCysGlyTyrArgArgCysArgAlaSerGlyValLeuThrThrSerMetGlyAsnThr
8430	AACTCGGGTACCGCAGGTGCGCGCAGCGCGTACTGACAACTAGTGTGGTAAACACC	8489	AACTCGGGTACCGCAGGTGCGCGCAGCGCGTACTGACAACTAGTGTGGTAAACACC

2735	QY	lleThrCysTyrValLysAlaLeuAlaCysLysAlaAlaGlylleAlaProThr	2754
8490	Db	CTCACTTGCTATCAATCAAGCCCGGCGACGCTGTCGAGCCGCGAGGCTCCAGGACTGCAC	8549
2755	QY	MetLeuValCysGlyAspLeuValLleSerGluSerGlnGlyThrGluGluAsp	2774
8550	Db	ATGCTCGTGTGGCGACGACTTAGTCGTATCTGTGAAGTGCGGGGTCCAGGAGGAC	8609
2775	QY	GluArgAsnLeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAsp	2794
8610	Db	GGCGGAGCCTGAGAGCCTTCACGGAGGCTATGACAGGACTCTCGCCCCCGGGGAC	8669
2795	QY	ProProArgProGluTyrAspLeuGluLeuLeuThrSerCysSerSerAsnValSerVal	2814
8670	Db	CCCCACAAACAGAAATACGACTTGGAGCTTATAACATCATGCTCTCCAAACGTGTCA	8729
2815	QY	AlaLeuGlyProGlnGlyArgArgTyrTyrLeuThrArgAspProThrThrProIle	2834
8730	Db	GGCCACGAGCGGCTGGAAAGAGGCTTACTACCTTACCCGTGACCTTACAAACCCCTC	8789
2835	QY	AlaArgAlaAlaTrrpGluThrValArgHisSerProValAsnSerTrpLeuGlyAsnIle	2854
8790	Db	GGGAGAGCGCGTGGGAGACACACACTCCAGTCAATCTCTGGTAGGCAACATA	8849
2855	QY	IleGlnTyrAlaProThrIleTrrpAlaArgMetValLeuMetThrHisPhePheSerIle	2874
8850	Db	ATCATGTTTGGCCCCACACTGTGGCGAGGATGATACTGATGACCAATTTCTTTAGCGTC	8909
2875	QY	LeuMetAlaGlnAspThrLeuAspGlnAsnLeuAsnPheGluMetTyrGlyAlaValTyr	2894
8910	Db	CTCATAGCAGCGGATACACTTGAACAGAGCTCTTAACCTGTGAGATCTACGGAGCCTGC	8969
2895	QY	SerValSerProLeuAspLeuProAlaIleIleGluArgLeuHisGlyLeuAspAlaPhe	2914
8970	Db	TCCATAGAACACATGGATCTACTCCATCAATCAATCAAGACTCCATGGCTCAGCGCATTT	9029
2915	QY	SerLeuHisThrTyrThrProHisGluLeuThrArgValAlaSerAlaLeuArgLysLeu	2934
9030	Db	TCACCTCCACAGTTACTCTCCAGGTGAATCAATAGGTTGGCGCATGCTCAGAAACATT	9089
2935	QY	GlyAlaProProLeuArgAlaTrrpLysSerArgAlaArgAlaValArgAlaSerLeuIle	2954
9090	Db	GGGGTCCCGCCCTTGGAGACTTGGAGACACCGGGCCCGAGCGTCCGCGCTAGGCTTCG	9149
2955	QY	SerArgGlyGlyArgAlaAlaValCysGlyArgTyrLeuPheAsnTrpAlaValLysThr	2974
9150	Db	TCCAGAGAGCGAGGCTGTATATGTGGCAAGTACCTCTTCAACTGGGAGTAGAACA	9209
2975	QY	LysLeuLysLeuThrProLeuProGluAlaArgLeuLeuAspLeuSerSerTrpPheThr	2994
9210	Db	AAGCTCAAACTCACTCCCAATAGCGCGCTGGCGGCTGGACTTGTCCGGTTGGTTACG	9269
2995	QY	ValGlyAlaGlyGlyAspIleTyrHisSerValSerArgAlaAArgProArgLeuLeu	3014
9270	Db	GCTGCTACAGCGGGGAGACATTATACAGCGTGTCTCATGCCCCCGCTGGTTC	9329
3015	QY	LeuPheGlyLeuLeuPheValGlyValGlyLeuPheLeuLeuProAlaArg	3033
9330	Db	TGGTTTTGCCCTACTCTCTCGCTGCAGGGGTAGGCATCTTACTCTCTCCCAACCGA	9386
RESULT 11			
ABK91431			
ID	ABK91431	standard; DNA; 9605 BP.	
XX	AC		
XX	AC	ABK91431;	
XX	AC		
XX	DT	15-NOV-2002 (first entry)	
XX	DE	Hepatitis C virus Con 1 isolate DNA mutant 8.	
XX	DE		
XX	KW	HCV; db; Con 1; adaptive mutation; liver failure; cirrhosis;	
XX	KW	hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;	
XX	KW	internal ribosome entry site; IRIS; NS5A; HCV replication; mutant.	

RESULT 11

ABK91431

ID ABK91431 standard; DNA; 9605 BP.

XX

AC ABK91431;

XX

DT 15-NOV-2002 (first entry)
yy

HE
XX
Hepatitis C virus Con 1 isolate DNA mutant A

DE HERPESVIRUS C VIRUS CON I ISOLATE DNA MUTANT 8:
XX

HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;

hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV; KW

KW internal ribosome entry site; IRES; NS5A; HCV replication; mutant.

6342	GCC	AAGCTC	ATGCCACAACTCGCTGGGATTCCTTTGTGTCTCTGCCAGCGCGGTATAGG	5401
2021	Gly	Val	TrpAlaGlyThrGlyIleMetThrThrArgCysProCysGlyAlaAsnIleSer	2040
6402	GGG	GCTCGGCGAGGAGACGGCATTTATGCACACTCGCTGCCACTGTGGAGCTGAGATCACT	6461	
2041	Gly	Asn	ValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTrp	2060
6462	GGACATGTCAAAACCGGACCATGAGGATCGTCGTCTCTAGGACCTGCAGGAACATGTGG	6521		
2061	Gln	GlyThrPheProIleAsnCysTyrThrGluGlyGlnCysValProLysProAlaPro	2080	
6522	AGTGGGAGCTTCCCATTAACGCCTACACACGGGCCCTGTACTCCCTCTCTGCGCGG	6581		
2081	Asn	PheLysValAlaIleTrpArgValAlaAlaSerGluTyrAlaGluValThrGlnHis	2100	
6582	AAC	TATAAGTTCCGCGCTGTCGAGGGTGTCTGCAGAGGAATACTGTGAGATGAAGCGGGTG	6641	
2101	Gly	SerTyrHisTyrIleThrGlyLeuThrThrAspAsnIleuLysValProCysGlnLeu	2120	
6642	GGG	GACTTCCAATACGTATCGGGTATGACTACTGTACAATCTTTAAATGCCCGTGCCAGATC	6701	
2121	Pro	SerProGluPhePheSerTrpValAspGlyValGlnIleHisArgPheAlaProThr	2140	
6702	CCATCGCCCCGAATTTTTCACAGATTGACGGGGTGGCGCTACACAGGTTGGGCCCCCT	6761		
2141	Pro	LysProPhePheArgAspGluValSerPheCysValGlyLysAsnSerPheValVal	2160	
6762	TGCAAGCCCTTGTGCGGGGAGGAGGTATCATTTACAGATGAGGACTCCACGAGTACCCGGTG	6821		
2161	Gly	SerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThr	2180	
6822	GGGTGCGCAATTTACCTTGGCGAGCCGGAACGGGACGTAGCCGTGTTGACGTCCTAATGCTCACT	6881		
2181	Asp	ProSerHisIleThrAlaGluThrAlaAlaArgArgLeuAlaArgGlySerProPro	2200	
6882	GATCCTCCCATATACAGACAGCGCGCGGAGAGGTTGGCGAGAGGTCACCCCT	6941		
2201	Ser	GluAlaSerSerAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThr	2220	
6942	TCTATGGCCAGCTCCTCGGCTAGCCAGCTGTCCGTCTCCATCTCTCAAGGCAACTTGCACC	7001		
2221	Thr	HisGlyLysAlaTyrAspValAspMetValAspAlaAsnLeuPhe	2236	
7002	GCCAAACCATACTCCCTCCGACCCGAGGCTCATAGAGGCTTAACCTCTCTGGAGGCGAGGAG	7061		
2237	Met	GlyGlyAspValThrArgIleGluSerGlySerLysValValValLeuAspSerLeu	2256	
7062	ATGGCGCGCAACATCACCAGGGTGTAGTCAGAGAACAAAGTGTGATTTGGACTCTTTC	7121		
2257	Asp	ProMetValGluGluArgSerAspLeuGluProSerIleProSerGluTyrMetLeu	2276	
7122	GATCCGCTTGTGGCAGAGGAGGATGAGGGGAGGTCTCCGTACCTCAGAAAATCTCTGCGG	7181		
2277	Pro	LysLysArgPheProProAlaLeuProAlaTrpAlaArgProAspTyrAsnProPro	2296	
7182	AAGTCTCGAGATTCCGCGGCGCTCGCCGCTGTGGCGCGCGCGAGCTACACACCCCGCG	7241		
2297	Leu	ValGluSerTrpLysArgProAspTyrGlnProAlaThrValAlaGlyCysAlaLeu	2316	
7242	CTAGTAGAGACGTGGAAGAAGCTGACTACGAACCACTCTGTGTGCTCATGGTGTCCCGCTA	7301		
2317	Pro	ProProArgLysThrProThrProProArgArgArgThrValGlyLeuSer	2336	
7302	CCACCTCCACGGTCCCTCTGTGCTCCGCTCGGAAAAGCGTACGGTGGTCTCACC	7361		
2337	Glu	AspSerIleGlyAspAlaLeuGlnLeuAlaIleLysSerPheGlyGlnProPro	2356	
7362	GAATCAACCTTACTTACTGCTTGGCCGAGCTTCCACCAAAAGTTTGGCAGCTCTCTCA	7421		
2357	Pro	SerGlyAspSerGlyLeuSerThrGlyAlaGlyAlaAlaAspSerGlySerGlnThr	2376	
7422	ACTTCGGCATTTACGGCGCAATACGACACATCTCTTGAGGCCCGCCCTTCTGGCTGC	7481		

QY 921 AlaLeuLeuArgMetCysThrMetAlaAlaArgHisLeuAlaGlyArgTyrValGlnMet 940
DB 3102 GGCCTTCTCGGATCTGGCGCTAGCGCGAAGATAGCGGAGGTCTATTAGGTGCAATG 3161
QY 941 AlaLeuLeuAlaLeuGlyArgTyrThrGlyThrIleTyrHisLeuThrProMet 960
DB 3162 GCCATCATCAAGTTAGGGCGCTTACTGGCAGCTATGTGTATACCATCTCAACCCCTCTT 3221
QY 961 SerAspTrpAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIleIlePhe 980
DB 3222 CGAGACTGGCGGCACAAACGGCTTCGAGATCTGGCGGTGGCTGTGAACAGTGTCTTC 3281
QY 981 SerProMetGluIleLysValIleValTrpGlyAlaGluThrAlaAlaCysGlyAspIle 1000
DB 3282 TCCCGAATGAGACCAAGCTCATCGTGGGGGCGAGATACCGCGCGTGGGTGACATC 3341
QY 1001 LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp 1020
DB 3342 ATCAACGGCTTGGCCGCTCTGCGCGTAGGGGCGAGAGATCTGCTTGGGGCGAGCGAC 3401
QY 1021 GlyTyrThrSerLysGlyTrpSerLeuLeuAlaProIleThrAlaValAlaGlnThr 1040
DB 3402 GGAATGCTCTCAAGGGGTGGAGGTGCTGGCGCCCATCACGGCGTACGCCACAGCAGCG 3461
QY 1041 ArgGlyLeuLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGlnAla 1060
DB 3462 AGAGGCTCTCTAGGTGTATAATCACCGCTGACTGGCGCGGACAAAACCAAGTGGAG 3521
QY 1061 GlyGluLeuGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly 1080
DB 3522 GGTGAGGTCCAGATCGTGTCACTGCTACCCAAACCTTCTGGCAACGTGCATCAATGG 3581
QY 1081 ValLeuTrpThrValThrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro 1100
DB 3582 GTATGCTGGAGTGTCTACACGGGGCGGAGCGAGGACCATCGCATCCACCAAGGGTCTCT 3641
QY 1101 ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTrpProSerProGly 1120
DB 3642 GTCATCCAGATGTATACCAATGTGGACCAAGACTTGTGGCTGGCGCGCTCTCAAGGT 3701
QY 1121 ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArgAsn 1140
DB 3702 TCCCGCTCATGACACCTGTACTCGGCTCTCGGACCTTTACCTGCTGTCAGAGGCAC 3761
QY 1141 AlaAspValIleProAlaArgArgGlyAspLysArgGlyAlaLeuLeuSerProArg 1160
DB 3762 GCCGATGTATTCCTCGCGCGCGCGAGGTGATAGCAGGGGTAGCTGTCTTGGCCCGG 3821
QY 1161 ProLeuSerThrLeuLysGlySerSerGlyProValLeuCysProArgGlyHisAla 1180
DB 3822 CCATTTCTTACTTGAAGGCTCTCGGGGGTCCGCTGTGTGCGCCCGGACACGCC 3881
QY 1181 ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle 1200
DB 3882 GTGGGCTATTACGGCGCGGTGTACCCCGTGGAGTGGCTAAAGCGTGGACTTTATC 3941
QY 1201 ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro 1220
DB 3942 CCTGTGGAGAACCTAGGGCAACCATGAGATCCCGGTGTTCACGGACCACTCTCTCCA 4001
QY 1221 ProAlaValProGlnThrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLys 1240
DB 4002 CCAGAGTGGCCAGAGCTTCAGGTGGCCCATCTGCATGCTCCCAACCGCGCGGTAAAG 4061
QY 1241 SerThrLysValProValAlaValAlaGlnGlyTyrLysValLeuValLeuAsnPro 1260
DB 4062 AGCACCAAGTCCCGCTGCTACGACGACCCAGCGGTACAAAGGTGTGTGTCTCAACCCC 4121
QY 1261 SerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerIleAlaHisGlyIleAsnPro 1280
DB 4122 TCTGTGCTGCAACGCTGGGCTTTGTGTCTTACATGTCCAAGGCCCATGGGTGTGATCT 4181

QY 1281 AsnIleArgThrGlyValArgThrValThrThrGlyAlaProIleThrTyrSerThrTyr 1300
DB 4182 AATATCAGGACCGGGGTGAGAACAAATACCACTGGCAGCCCATACGCTACTCCACCTAC 4241
QY 1301 GlyLysPheLeuAlaAspGlyCysAlaGlyGlyAlaTyrAspIleIleIleCysAsp 1320
DB 4242 GGCAGAGTTCTTGGCGAGCGCGGTGCTCAGAGGTGCTTATGCACATAATAATTTGTGAC 4301
QY 1321 GluCysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGlnAla 1340
DB 4302 GAGTGCCACTCCACGATGCCATCCATCTTGGGCATCGGCATCTCTTGCACCAAGCA 4361
QY 1341 GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProGlySerValThr 1360
DB 4362 GAGACTGGGGGCGAGAGCTGTGTGTCTCGCATCTGCTACCCCTCCGGGCTCCGCTCACT 4421
QY 1361 ThrProHisProAsnIleGluGluValAlaLeuGlyGlnGluGlyGluIleProPheTyr 1380
DB 4422 GTGTCCCATCTTAACATCGAGAGGTGTCTGTCTCCACCACCGAGAGATCCCTTTTAC 4481
QY 1381 GlyArgAlaIleProLeuSerTyrIleLysGlyArgHisLeuIlePheCysHisSer 1400
DB 4482 GGCAGGCTATCCCCCTCGAGGTGATCAAGGGGGGAGACATCTCATCTTCTGCCACTCA 4541
QY 1401 LysLysLysCysAspGluLeuAlaAlaLeuArgGlyMetGlyLeuAsnSerValAla 1420
DB 4542 AAGAAGAGTGCAGCAGCTCGCGGAAGCTGGTCGATTCGATTCGATTCGATTCGATTCG 4601
QY 1421 TyrTyrArgGlyLeuAspValSerValIleProThrGlnGlyAspValValValAla 1440
DB 4602 TACTACCGGCTTGTGAGCTGTCTCATCCGACCGCGGCGATGTGTCTGCTGCTGCTG 4661
QY 1441 ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnVal 1460
DB 4662 ACCGATGCTCTCATGCTGCTTACCGCGACTTCGACTCTGTGATGACTGCACACG 4721
QY 1461 AlaValThrGlnValValAspPheSerLeuAspProThrPheThrIleThrGlnIle 1480
DB 4722 TGTGTCACTCAGACAGTCGATTCAGCTTGACCTTGACCTTACCATTCGAGACACACG 4781
QY 1481 ValProGlnAspAlaValSerArgSerGlnArgArgGlyArgThrGlyArgGlyLeu 1500
DB 4782 CTCCTCCAGGATGCTCTCCAGGACTCAACCGCGGGGCGAGCTGGCAGGGGAGGCCA 4841
QY 1501 GlyIleTyrArgTyrValSerThrGlyGluArgAlaSerGlyMetPheAspSerValVal 1520
DB 4842 GGCATCTATAGATTTGTGGCAGCGGGGAGCGCCCTCCGGCATGTTCGACTGCTCCGCTC 4901
QY 1521 LeuCysGluCysTyrAspAlaGlyAlaAlaTyrTyrGluLeuThrProSerGluThrThr 1540
DB 4902 CTCGTGAGTGTATGACCGGGGCTGTCTTGTGTATGAGTCAACGCCCGCGAGACTACA 4961
QY 1541 ValArgLeuArgAlaTyrPheAsnThrProGlyLeuProValCysGlnAspHisLeuGlu 1560
DB 4962 GTTAGGTACGAGCGTACATGAACACCCCGGGGCTTCCGCTGTCGAGAGCACCATTCTGA 5021
QY 1561 PheTrpGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThr 1580
DB 5022 TTTTGGAGGGGCTCTTACGGGCTCTCATATAGATGCCCTATTTATCCCCAGACA 5081
QY 1581 LysGlnSerGlyGluAsnPheAlaTyrLeuThrAlaTyrGlnAlaThrValCysAlaArg 1600
DB 5082 AAGCAGAGTGGGAGAACTTTCTTACCTGGTAGGTACCAAGCCACCGGTGGCTAGG 5141
QY 1601 AlaLysAlaProProSerTrpAspValMetTrpLysCysLeuThrArgLeuLysPro 1620
DB 5142 GCTCAAGCCCTCCCGCATCGTGGGACAGATGTGGAAGTGTGTGATCGCGCTTAAACCC 5201
QY 1621 ThrLeuValGlyProThrProLeuLeuTyrArgLeuGlySerValThrAsnGluValThr 1640
DB 5202 ACCCTCCATGGGCCAACCCCTGTCTATACAGCTGGGGCTGTTCAGATGAAGTCAACC 5261
QY 1641 LeuThrHisProValThrLysTyrIleAlaThrCysMetGlnAlaAspLeuGluValMet 1660

Db	882	CTGCTGCTGTCATCACCCCGGCTCTCGCTGCGAGTGAAGAACATCAGTACCGGC	941	1962	TTGAACAGCAGCCTCGACACCGCTGGGGTCTATGTTGGGCTGCACGTGAGTGAACCTCTTCT	2021
Qy	201	TyrMetValThrAsnAspCysThrAsnAspSerIleThrTrpGlnLeuGlnAlaVal	220	561	GlyTyrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer	580
Db	942	TACATGGTGAATACAGCATGACCAATGACAGCATTTACCTGGCAGCTCCAGGCTGCTGC	1001	2022	GGCTACACAGACTTGGCGGCACACCTGCGCTACTAGAGCTGACTTCAACGCCAGC	2081
Qy	221	LeuHisValProGlyCysValProCysGlyLysValGlyAsnAlaSerGlnCysTrpIle	240	581	ThrAspLeuLeuCysProThrAspCysPheArgLysHisProAspThrThrTyrLeuLys	600
Db	1002	CTCACGTCCTCCGGTGGCTCCGTCGCGAAGAGTGGGAGTGCATCTCAGTCTGGATA	1061	2082	ACCGACCTGTTGTCGCCACCGACTGTTTGTAGAAAGCATCTCTGATACCACTTACCTCAA	2141
Qy	241	ProValSerProHsnValAlaValGlnArgProGlyAlaLeuThrGlnGlyLeuAsgThr	260	601	CysGlySerGlyProTrpLeuThrProArgCysLeuIleAspTyrProTyrArgLeuTrp	620
Db	1062	CCGCTCTCACCGAATGTGGCGTGCAGCGCCGGCGCTCTACGAGGCTTGGCGAGC	1121	2142	TGCGGCTCTGGCGCTGGCTCAGCGCAAGTGCCTGATCGACTACCCCTACAGGCTCTCG	2201
Qy	261	HisIleAspMetValValMetSerAlaThrIleuCysSerAlaLeuTyrValGlyAspLeu	280	621	HisTyrProCysThrValAsnTyrThrIlePheLysIleArgMetTyrValGlyGlyVal	640
Db	1122	CACATCGACATGGTGTGTGATGTCGCGCACGCTCTGCTCTGCTCTACGTTGGGGACCTC	1181	2202	CATTACCCCTGCACAGTAACTATACCATCTTCAAAATAAGGATGTATGTGGGAGGGTT	2261
Qy	281	CysGlyGlyValMetLeuAlaGlnMetPheIleValSerProGlnHisIleTyrPhe	300	641	GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp	660
Db	1182	TGCGGTGGGGTGGTCTCGAGCCCAATGTTTCAATGCTCTGCGCGCAGCACCACTGGTTT	1241	2262	GAGCACAGGCTCAGCGCTGCATGCAATTTCTCTGCTGGGGATCGTTGCAACTTGGAGGAC	2321
Qy	301	ValGlnAspCysAsnCysSerIleTyrProGlyThrIleThrGlyHisArgMetAlaTrp	320	661	ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGluTrpAlaIleLeuPro	680
Db	1242	GTCCAGACTGCAATGCTCCATCTACCTGTTGATCATCTGACGACACCGCATGGCATGG	1301	2322	AGAGACAGAGTCAACTGTCTCTTGTGTCATCCACACCGAATGGGCCATTTTACCT	2381
Qy	321	AspMetMetMetAsnTrpSerProThrAlaThrMetIleLeuAlaTyrAlaMetArgVal	340	681	CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisIleGlnAsnIle	700
Db	1302	GACATGATGATGAATGCTGCCCGCACCGCTACCATGATCTTGGCGTACGGGATGCGTGC	1361	2382	TGCTCTTACTCGGACCTGCCGCTTGTGCTGCTGCTCTCTCACCCTCCACCAACATC	2441
Qy	341	ProGluValIleIleAspIleIleSerGlyAlaHisIleTrpGlyValMetPheGlyLeuAla	360	701	ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrIleTyrIleValArgTrp	720
Db	1362	CCCAGGTCAATATAGACATCATTTAGCGGGCTCATTTGGGGCTCATGTTGGCTTGGCC	1421	2442	GTGACGTCAATTCATGTATGCGCTATCACCCTGCCCTCACAATAATACATCGTCCGATG	2501
Qy	361	TyrPheSerMetGlnGlyAlaTrpAlaLysValValIleLeuLeuLeuAlaGly	380	721	GluTrpValIleLeuLeuPheLeuLeuAlaAspAlaArgValCysAlaCysLeuTrp	740
Db	1422	TACTTCTCTATGAGGAGGCTGGCGGAAAGTCTGTTGTCATCTCTTGTGGCGCGCGG	1481	2502	GAGTGGGTAACTACTTATTTCTGCTTTAGGCGACCGCAGGGTTTGGCGCTCTATGG	2561
Qy	381	ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr	400	741	MetLeuIleLeuLeuGlyGlnAlaGluAlaLeuGluLysLeuValIleLeuHisAla	760
Db	1482	GTGACGCGCGCACCATCTACTGTTGGGGGTTCTCGCGCAGACACCGCGCGCTCACC	1541	2562	ATGCTCATCTTGTGGCGCAGCGCGGAGGAGCTTTGGAGAACCTCGTAATCTCAATGCA	2621
Qy	401	SerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTrp	420	761	AlaSerAlaAlaSerCysAsnGlyPheLeuTyrPheValIlePhePheValAlaAlaTrp	780
Db	1542	AGCTTATTTGACATGGGCCCCGAGCAGAAATCCAGCTCGTTTACCAATGGCAGCTGG	1601	2622	GCATCCCTGGCGCGCAGCAGCTCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2681
Qy	421	HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer	440	781	TyrIleLysGlyArgValValProLeuAlaThrTyrSerLeuThrGlyLeuTrpSerPhe	800
Db	1602	CACATCAACCGCACCGCCCTGAACTGCAATGACTCTCTTGACACCGGCTTTATGCGTCT	1661	2682	TATCTGAAGGTAGGTGGGTGGCGGAGCGGCTTACGCGCTCTACCGGATGTGCGCTCTC	2741
Qy	441	LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg	460	801	SerLeuLeuLeuLeuAlaLeuProGlnGlnAlaTyrAlaTyrAspAlaSerValHisGly	820
Db	1662	CTGTTCTACACCCACAGCTTCAACTCTGTCAGGATGTCCGAAACGATGTCCGCTGCCGC	1721	2742	CT	2801
Qy	461	SerIleGluAlaPheArgValGlyTrpGlyAlaLeuGlnTyrGluAspAsnValThrAsn	480	821	GlnIleGlyAlaAlaLeuLeuValMetIleThrLeuPheThrLeuThrProGlyTyrLys	840
Db	1722	AGTATCGAGGCTCTCCGGGTGGAGTGGGGCGCTTTCATATGAGGATAATGTCAACAT	1781	2802	TCGTGTGGCGGCTGTTCTTGTGGGTAAATGGCGCTGACTCTGTGCGCATATTACAAG	2861
Qy	481	ProGluAspMetArgProTyrCysTrpHisTyrProArgGlnCysGlyValValSer	500	841	ThrLeuLeuSerArgPheLeuTrpLeuCysTyrLeuLeuThrLeuGlyGluAlaMet	860
Db	1782	CCAGAGGATATGAGACCTTATGTGGCACTACCCACCAAGGAGTGTGGGCTGTCTCC	1841	2862	CGCTATATACGTGTGTCATGTGGTCTTCAATTTCTGACAGAGTAGAGAGGAGCAA	2921
Qy	501	AlaLysThrValCysGlyProValTyrCysPheThrProSerProValValGlyThr	520	861	ValGlnGluTrpAlaProProMetGlnValArgGlyGlyArgAspGlyIleIleTrpAla	880
Db	1842	GCGAAGACTGTGTGGGCCAGTGTACTGTTTCAACCCCGAGGAGTGTAGTGGGACG	1901	2922	CTGACGTGTGGGTCTCCCGCTCAACGTCGCGGGGGGGCGGATGCGGCTATCTTACTC	2981
Qy	521	ThrAspArgLeuGlyAlaProThrTyrThrTrpGlyGluAsnGluThrAspValPheLeu	540	881	ValAlaIlePheTyrProGlyValValPheAspIleThrIleThrTrpLeuLeuAlaValLeu	900
Db	1902	ACCGACAGCTGTGAGCGCCCATTTACACGTGGGGGAGAGATGACAGATGTCTTCTTA	1961	2982	ATGTGTGTAGTACACCGACCCCTGGTATTGACATCACCAAACTACTCTCTGGCCATCTTC	3041
Qy	541	LeuAsnSerThrArgProProLeuGlySerTrpPheGlyCysThrTrpMetAsnSerSer	560	901	GlyProAlaTyrLeuLeuLysGlyAlaLeuThrArgValProTyrPheValArgAlaHis	920
				3042	GGACCCCTTGGATTTCTTCAAGCCAGTGTGTTGTTTAAAGTCCCTTACTCTGCGCGGTTCAA	3101

QY 2835 AlaArgAlaAlaTrpGluThrValArgHisSerProValAsnSerTrpLeuGlyAsnIle 2854
 DB 8790 GCGAGAGCGCGTGGGAGACAGACACACACACACACACACACACACACACACAT 8849
 QY 2855 IleGlnTyrAlaProThrIleTrpAlaArgMetValLeuMetThrHisPheSerIle 2874
 DB 8850 ATCATGTGTTCCCGCCACACTGTGGCGAGGATGATGATGATGATGATGATGATG 8909
 QY 2875 LeuMetAlaGlnAspThrLeuAspGlnAsnLeuAsnPheGluMetTyrGlyAlaValTyr 2894
 DB 8910 CTATAGCCAGGATCAGCTTGACAGGCTTTAACTGTGAGATCTACGAGGCTCTAC 8969
 QY 2895 SerValSerProLeuAspLeuProAlaIleGluArgLeuHisGlyLeuAspAlaPhe 2914
 DB 8970 TCCATAGAACCACTGGATCTACCTCCAAATCAATCAAGAGCTCCATGCGCTCAG 9029
 QY 2915 SerLeuHisThrTyrThrProHisGluLeuThrArgValAlaSerAlaLeuArgLysLeu 2934
 DB 9030 TCACTCCACAGTTACTCTCCAGGTGAATCAATAGGGTGGCCGCTCAGAAACTT 9089
 QY 2935 GlyAlaProProLeuArgAlaTrpLysSerArgAlaArgAlaValArgAlaSerLeuIle 2954
 DB 9090 GGGTCCCGCCCTTGGAGCTTGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCG 9149
 QY 2955 SerArgGlyGlyArgAlaAlaValCysGlyArgTyrLeuPheAsnTrpAlaValLysThr 2974
 DB 9150 TCCAGAGGAGCGAGGCTGTATATGTGGCAAGTACTCTTCACTGGCGAGTAAGAAC 9209
 QY 2975 LysLeuLysLeuThrProLeuProGluAlaArgLeuLeuAspLeuSerTrpPheThr 2994
 DB 9210 AAGCTCAAACTCACTCCAAATAGCGCGCGCTGGCGGCTGGACTGTCCGGTTGGTTCAG 9269
 QY 2995 ValGlyAlaGlyGlyAspIleTyrHisSerValSerArgAlaArgProArgLeuLeu 3014
 DB 9270 GCTGGCTACAGCGGGGAGACATTTATCAGCGTGTCTCATGCGCGCGCGCTGGTTC 9329
 QY 3015 LeuPheGlyLeuLeuLeuPheValGlyValGlyLeuPheLeuLeuProAlaArg 3033
 DB 9330 TGGTTTGGCTACTCTCTGCTGCTGCGAGGGTAGGAGATCTACTCTCTCCCAACCGA 9386
 RESULT 10
 AAC86648
 ID AAC86648 standard; DNA; 9611 BP.
 XX AAC86648;
 AC
 XX
 DT 02-APR-2001 (first entry)
 XX
 DE Nucleotide sequence of chimeric Hepatitis C virus clone pH77(p7)-J6S.
 XX HCV; HCV strain HC-J6CH; HCV genotype 2a; antiviral; vaccine; ss.
 OS Synthetic.
 OS Hepatitis C virus.
 XX
 XX
 FT Key Location/Qualifiers
 FT CDS 342..9389
 FT /*tag= a
 XX
 XX WO200075338-A2.
 XX
 XX
 PD 14-DEC-2000.
 XX
 PF 02-JUN-2000; 2000WO-US15446.
 XX
 XX 04-JUN-1999; 99US-0137693.
 XX
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 XX Yanagi M, Bukh J, Emerson SU, Purcell RH;
 XX WPI; 2001-061728/07.
 DR

P-PSDB; AAB30733.

XX Nucleic acid molecule encoding human hepatitis C virus of genotype 2a
 PT for developing vaccines, for diagnosis of hepatitis C virus and in
 PT screening assays for identification of antiviral agents -
 XX
 PS Disclosure; Page 146-149; 167pp; English.

XX AAC86645-48 represent chimeric cDNA clones of infectious Hepatitis C
 CC virus (HCV). In each clone the C, E1 and E2 genes are derived from HCV
 CC strain HC-J6CH genotype 2a. The p7 protein was derived from HCV strain
 CC HC-J6CH or HCV strain pCV-H77C, and the NS genes were derived from HCV
 CC strain pCV-H77C. Such HCV sequences are capable of expressing the virus
 CC when transfected into cells. The HCV protein is useful for assaying
 CC candidate antiviral agents for activity against HCV. Antibodies specific
 CC for HCV polypeptide are useful in prevention and treatment of diseases
 CC caused by HCV in animals, in particular humans. The HCV polypeptides
 CC serve as immunogens in the development of vaccines for preventing HCV
 CC in mammals or as antigens in diagnostic assays for detecting the
 CC presence of HCV in biological samples. The HCV polynucleotide is also
 CC useful for identifying cell lines capable of supporting the replication
 CC of HCV in vitro and to produce attenuated viral strains via passage in
 CC vitro or in vivo.

SQ Sequence 9611 BP; 1925 A; 2897 C; 2695 G; 2094 T; 0 other;

Alignment Scores:
 Pred. No.: 0 Length: 9611
 Score: 12925.00 Matches: 2371
 Percent Similarity: 86.90% Conservative: 270
 Best Local Similarity: 78.02% Mismatches: 368
 Query Match: 80.18% Indels: 30
 DB: 22 Gaps: 4

US-09-980-559-2 (1-3033) x AAC86648 (1-9611)

QY 1 MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGln 20
 DB 342 ATGAGACACAAATCTTAACCTCAAGAGAAACCAAGAAACACCAACCGTCCGCCA 401
 QY 21 AspValLysPheProGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArg 40
 DB 402 GACGTTAAGTTTCCGGCGCGCGCCAGATCGTTGGCGAGTATATTGTTGGCGGAG 461
 QY 41 GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly 60
 DB 462 GCGCCAGGTTGGGTGGCGCGCGCAAGAGAACTTCGAGCGGTCCAGCACCGTGA 521
 QY 61 ArgArgGlnProIleProLysAspArgArgSerThrGlyLysSerTrpGlyLysProGly 80
 DB 522 AGCGCCAGCCCATCCCTAAGATCGCGCTCCACTGGCAATCTGGGGAAACACGGA 581
 QY 81 TyrProTrpProLeuTyrGlyAsnGluGlyLeuGlyTrpAlaGlyTrpLeuLeuSerPro 100
 DB 582 TACCCCTGGCCCTTATACGGAAATGAGGAGTCTGGCTGGCAGGATGGCTCTGTCCCC 641
 QY 101 ArgGlySerArgProSerTrpGlyProAsnAspProArgHisArgSerArgAsnValGly 120
 DB 642 CGAGGTTCCCGTCCCTCTTGGGGCCCCCAATGACCCCCGGCATAGGTCCGCAACGTGG 701
 QY 121 LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProValVal 140
 DB 702 AAGTCATCGATACCTTAACGTGGCGCTTTGCCGACCTCATGGGGTACATCCCTGTCTG 761
 QY 141 GlyAlaProLeuGlyGlyValAlaArgAlaLeuAlaHisGlyValArgValLeuGluAsp 160
 DB 762 GCGCGCCGCTCGCGCGGTCCAGAGCTCTCGCGCATGGCGTGGAGAGTCTCTGAGGAC 821
 QY 161 GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla 180
 DB 822 GGGGTAAATTTTGCAACAGGAACTTACCCGGTGGCTCTCTTCTATCTTCTGTGGCC 881
 QY 181 LeuLeuSerCysIleThrThrProValSerAlaAlaGluValLysAsnIleSerThrGly 200

Db	6642	: : :
----	------	--

Db	2262	GAGCAGCGCTCAGCGCTGCATCAATTTCTACCTCGTGGGATCGTTGCAACTTTGGAGGAC	2321
Qy	661	ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGlnTrpAlaIleLeuPro	680
Db	2322	AGAGACAGAGTCAACTGCTCTCTTTGTTGCACTCCACCAACGGAATGGGCAATTTACCT	2381
Qy	681	CysSerTySerAspLeuProAlaLeuSerThrGlyLeuLeuHisGlnAsnIle	700
Db	2382	TGCTCTTACTCGGACCTGCCCGCTGTGCACTGGTCTTCTCCACTCCACCAAAACATC	2441
Qy	701	ValAspValGlnPheMetTyGlyLeuSerProAlaLeuThrbyTyIleValA-gTrp	720
Db	2442	GTGAGAGTCAATTCATGTATGCTTATCACTGCCCTCACAATAATACATCGTCCGATGG	2501
Qy	721	GlnTrpValIleLeuLeuPheLeuLeuAlaAspAlaArgValCysAlaCysLeuTrp	740
Db	2502	GAGTGGGTAATACTCTTATTTCTGCTTTAGCGAGCCAGGCTTTGCGCTCTCTATGG	2561
Qy	741	MetLeuIleLeuLeuGlyGlnAlaGluAlaLeuGluLysLeuValIleLeuHisAla	760
Db	2562	ATGCTCATCTTTGGGCGAGCGGAGAGCTTTGGAGAACCTCGTAATCTCAATGCA	2621
Qy	761	AlaSerAlaAlaSerCysAsnGlyPheLeuTy-PheValIlePhePheValAlaAlaTrp	780
Db	2622	GCATCCCTGCGCGGACGCGCTTTGCTGCTCTCTGCTGCTCTCTGCTGCTGCTGCTG	2681
Qy	781	TyrIleLysGlyArgValProLeuAlaThrTySerLeuThrGlyLeuTrpSerPhe	800
Db	2682	TATCTGAAGGTTAGTGGTGGCGCGAGCGGCTCTACGCGCTCTACGCGGATGTGCGCTCTC	2741
Qy	801	SerLeuLeuLeuAlaLeuProGlnGlnAlaTyAlaTyAlaTyAlaSerValHisGly	820
Db	2742	CTCTGCTCTCTGCTGGGCTTCCCTCAGCGGCATATGCACTGGACGGAGTGGCGCG	2801
Qy	821	GlnIleGlyAlaAlaLeuLeuValMetIleThrLeuPheThrLeuThrProGlyTyLys	840
Db	2802	TCGTGTGGCGCGTGTCTTGTGGGTAAATGGCGCTGCTCTGTCGCCATATTACAG	2861
Qy	841	ThrLeuLeuSerArgPheLeuTrpTrpLeuCysTyLeuLeuThrLeuGlyGluAlaMet	860
Db	2862	CGCTATATACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2921
Qy	861	ValGlnGlnTrpAlaProMetGlnValArgGlyArgAspGlyIleIleTrpAla	880
Db	2922	CTGACGCTGGTGTCTCCCTCCCTCAAGCTCGGGGGGGGGCGGATGCGGTCATCTTACT	2981
Qy	881	ValAlaIlePheTyProGlyValValPheAspIleThrLysTrpLeuLeuAlaValLeu	900
Db	2982	ATGTGTGTAGTACACCGCCCTGTTATTTGACATCACCACCACTCTCTGGCCATCTTC	3041
Qy	901	GlyProAlaTyLeuLeuLysGlyAlaLeuThrArgValProTyPheValArgAlaHis	920
Db	3042	GGACCCCTTTGGATCTTCAAGCCAGTGTGCTTAAAGTCCCTACTTCTGCGCGCTTCAA	3101
Qy	921	AlaLeuLeuArgMetCysThrMetAlaAlaGhiSLeuAlaGlyArgTyValGlnMet	940
Db	3102	GGCCTTCTCCGGATCTCGCGCTAGCGGAGATAGCCGAGGTCTATTAGTGCATATG	3161
Qy	941	AlaLeuLeuAlaLeuLeuGlyArgTrpThrGlyThrTyIleTyAspHisLeuThrProMet	960
Db	3162	GCCATCATCAAGTTAGGGCGCTTACTGGCACCTATGTATATAACCATCTCACCCCTCT	3221
Qy	961	SerAspTrpAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIleIlePhe	980
Db	3222	CGAGACTGGGCGCACACCGGCTCGAGATCGCGCTGGGTGGGAACCACTCGCTCTTC	3281
Qy	981	SerProMetGluLysValIleValTrpGlyAlaGluThrAlaAlaCysGlyAspIle	1000
Db	3282	TCCGAATGGAGACCAAGCTCATACGTGGGGGGGAGATACCGCGCGGTGCGGTGACATC	3341
Qy	1001	LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp	1020
Db		GAGTCCCATCTTAACATCGAGGAGTGTCTCTGTGCCACCCGAGAGATGCCCTTTTAC	4481

CC presence of HCV in biological samples. The HCV polynucleotide is also
 CC useful for identifying cell lines capable of supporting the replication
 CC of HCV in vitro and to produce attenuated viral strains via passage in
 CC vitro or in vivo.
 XX
 SQ Sequence 9611 BP; 1924 A; 2898 C; 2694 G; 2095 T; 0 other;

Alignment Scores:

Pred. No.: 0 Length: 9611
 Score: 12925.00 Matches: 2371
 Percent Similarity: 86.90% Conservative: 270
 Best Local Similarity: 78.02% Mismatches: 368
 Query Match: 80.18% Indels: 30
 DB: 22 Gaps: 4

US-09-980-559-2 (1-3033) x AAC86646 (1-9611)

QY 1 MetSerThrAsnProIysProGlnArgLysThrLysArgAsnThrAsnArgArgProGln 20
 DB 342 ATGAGCACAATCTTAACCTCAAGAAAAACCAAGAAACACCAACACCGTCCGCCACAA 401
 QY 21 AspValLysPheProGlyGlyGlnIleValGlyGlyValTyrlLeuLeuProArgArg 40
 DB 402 GAGCTTAAGTTTCCGGCGCGCCAGATCGTTGGCGAGTATACCTTGTTCGCGCAGG 461
 QY 41 GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly 60
 DB 462 GGCCTCCAGGTGGGTGGCGCGGACNAGNAGNCTTCGGAGCGGTCCACGACACCTGGA 521
 QY 61 ArgArgGlnProIleProIysAspArgArgSerThrGlyLysSerTrpGlyLysProGly 80
 DB 522 AGCGCGCAGCCATCCCTAAGATCGCGCTCCACTGGCAATCCTCTGGGAAAAACAGGA 581
 QY 81 TyrProTrpProLeuTyrlAsnGluGlyLeuGlyTyrlAlaGlyTyrlLeuLeuSerPro 100
 DB 582 TACCTCCCTGGCCCTATACGCGGAATGAGGGACTCGCTGGCGAGGATGGCTCTGTCTCCC 641
 QY 101 ArgGlySerArgProSerTrpGlyProAsnAspProArgHisArgSerArgAsnValGly 120
 DB 642 CGAGGTTCCTCCCTCTTGGGGCCCAATAGACCCCGCGCATAGGTTCGGCAACGTGGGT 701
 QY 121 LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrlProValVal 140
 DB 702 AAGGTCATCATACCTAACGTGCGCTTTGGCGACCTCATGGGGTACATCCCTGTCTGTG 761
 QY 141 GlyAlaProLeuGlyGlyValAlaArgAlaLeuAlaHisGlyValArgValLeuGluAsp 160
 DB 762 GCGCGCCCGCTCGCGCGCGTCGCGAGCTCTCGCGCATGGCGTGAGAGTCTCTGGAGGAC 821
 QY 161 GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerilePheLeuLeuAla 180
 DB 822 GGGGTTAATTTGCAACAGGAACTTACCGGTGTCTCTTTCTATCTTCTGTCTGGCC 881
 QY 181 LeuLeuSerCysIleThrProValSerAlaAlaGluValLysAsnIleSerThrGly 200
 DB 882 CTGCTGTCTGCATCACCACCCCGCTCTCGCTGCGAAGTGAAGAACATCAGTACCGGC 941
 QY 201 TyrMetValThrAsnAspCysThrAsnAspSerileThrTrpGlnLeuGlnAlaVal 220
 DB 942 TACATGGTGACTAACCACTGACCAATGACAGCATTTACCTGGCAGCTCCAGGCTGTCTGC 1001
 QY 221 LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTrpIle 240
 DB 1002 CTCACGTCCTCCGGGTGGTCCCGTGGAGAAAGTGGGAATGCATCTCAGTGTCTGGATA 1061
 QY 241 ProValSerProAsnValAlaValGlnArgProGlyAlaLeuThrGlnGlyLeuArgThr 260
 DB 1062 CCGGTCTCACCGAATGTGGCGGTGCAGCGCCCGGCGCTCACGCGAGGCTTGGCGAGC 1121
 QY 261 HisIleAspMetValValMetSerAlaThrLeuCysSerAlaLeuTyrlValGlyAspLeu 280
 DB 1122 CACATCGCATGTTGTGTGTCCTCCGCGACGCTCTGCTCTGCTCTACGCTGGGGACCTC 1181

QY 281 CysGlyGlyValMetLeuAlaAlaGlnMetPheIleValSerProGlnHisIleTrpPhe 300
 DB 1182 TCGGGTGGGGTGATGCTCGCAGCCCAAAATGTTTCAATGTCTCGCGCAGCACCACTGGTTT 1241
 QY 301 ValGlnAspCysAsnCysSerileTyrlProGlyThrIleThrGlyHisArgMetAlaTrp 320
 DB 1242 GTCCAAGACTGCAATTTGCTTCCATCTACCTCGTGTACCATCTGACACCGCATGGCATGG 1301
 QY 321 AspMetMetMetAsnTrpSerProThrAlaThrMetIleLeuAlaTyrlAlaMetArgVal 340
 DB 1302 GACATGATGATGAATGTCGCGCCACGGCTACCATGATCTTGGCGTACCGATGCGGTGTC 1361
 QY 341 ProGluValIleIleAspIleIleSerGlyAlaHisTrpGlyValMetPheGlyLeuAla 360
 DB 1362 CCGGAGGTCAATATAGACATCATTAGCGGGCTCATTTGGGGCTCATGTTGGCTTGGCC 1421
 QY 361 TyrPheSerMetGlnGlyAlaTrpAlaLysValValIleLeuLeuLeuAlaGly 380
 DB 1422 TACTTCTCTATGACGGAGCGTGGCGCAAGTGTGTGTCTCTCTTGTGGCGCGCGG 1481
 QY 381 ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr 400
 DB 1482 GTGGAGCGCGCACCATCTGTTGGGGTCTTGGCGCGACACACCGGGCGCTTCACC 1541
 QY 401 SerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTrp 420
 DB 1542 AGCTTATTGACATGGGCCCGCAGGAGCAAAATCCAGCTCGTTAAACCAATGGCAGCTGG 1601
 QY 421 HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer 440
 DB 1602 CACATCAACCGCACCGCTGAATGCAATGACTCTCTTGACACACCGGCTTATTCGGTCT 1661
 QY 441 LeuPheTyrlThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg 460
 DB 1662 CTGTCTTACACCCAGCTTCAACTCGTCAGATGTCCGAGACGATGTCGCGCTGCCGC 1721
 QY 461 SerIleGluAlaPheArgValGlyTrpGlyAlaLeuGlnTrpGluAspAsnValThrAsn 480
 DB 1722 AGTATCGAGGCTTTCGGGGTGGGATGGGGCGCTTGCATATATGAGGATATGTACCANT 1781
 QY 481 ProGluAspMetArgProTyrlCysTrpHisTyrlProProArgGlnCysGlyValValSer 500
 DB 1782 CCAGAGGATATGAGACCTTATTTGCTGGCACTACCCACCAAGGAGTGTGGGTCTCC 1841
 QY 501 AlaLysThrValCysGlyProValTyrlCysPheThrProSerProValValValGlyThr 520
 DB 1842 GCGAAGACTGTGTGGCGCAGTACTGTTTACCCCGCCAGCTGTGTGTGTGGGCAGC 1901
 QY 521 ThrAspArgLeuGlyAlaProThrTyrlThrTrpGlyGluAsnGluThrAspValPheLeu 540
 DB 1902 ACCGACAGCTTGGAGCGCCCACTTACAGTGGGGGAGAAATGAGACAGATGTCTTCTTA 1961
 QY 541 LeuAsnSerThrArgProProLeuGlySerTrpPheGlyCysThrTrpMetAsnSerSer 560
 DB 1962 TTGAACAGCACTCGACCCCGCTGGGTGTGTTGGGTGTGCGCTGCGTGTGAGTGAACCTTCT 2021
 QY 561 GlyTyrlThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer 580
 DB 2022 GGTACACCAAGACTTGGCGGCGACACCTTGGCGGTACTAGAGCTGACTTCAACCGCAGC 2081
 QY 581 ThrAspLeuLeuCysProThrAspCysPheArgLysHisProAspThrThrTyrlLeuLys 600
 DB 2082 ACGGACCTTGTGTGCGCCCGACGACTGTTTAGGAGAGCATCTGTATACCACTTACCTCAA 2141
 QY 601 CysGlySerGlyProTrpLeuThrProArgCysLeuIleAspTyrlProTyrlArgLeuTrp 620
 DB 2142 TCGCGCTCTGGCGCTCGCTACGCAAGGTGCTGATCGACTACCCCTTACAGGCTCTGG 2201
 QY 621 HisTyrlProCysThrValAsnTyrlThrIlePheLysIleArgMetTyrlValGlyVal 640
 DB 2202 CATTACCCCTGACAGTAACTATACCATCTTCAAAATAAGGATGTATGTGGAGGGGT 2261
 QY 641 GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp 660

9030	TCATCCACAGTTACTCTCCAGGTGAATCAATAGGTTGGCCGCATGCTCTCAGAAAACCTT	9088
2935	GlyAlaProLeuArgAlaTrpIysSerArgAlaArgAlaValArgAlaSerLeuIle	2954
9090	GGGGTCCCGCCCTTGCAGAGCTTGGAGACACCGGGCCGAGGCTCCGGCTAGGCTTCGTG	9149
2955	SerArgGlyGlyArgAlaAlaValCysGlyArgTyrLeuPheAsnTrpAlaValIysThr	2974
9150	TCCAGAGAGAGCAGGGCTGCTATATGTGGCAAGTACCTCTTCAACTGGGCAGTAGAACA	9209
2975	LysLeuIysLeuThrProLeuProGluAlaArgLeuLeuAspLeuSerSerTrpPheThr	2994
9210	AAGCTCAAACTCACTCCAAATAGCGCCCGCTGCGCGGCTGGACTTGTCCGGTTGGTTCA	9269
2995	ValGlyAlaGlyGlyGlyAspIleTyrHisSerValSerArgAlaArgProArgLeuLeu	3014
9270	GCTGGCTACAGCGGGGAGACATTATCAGACGGTGTCTCATGCGCGGCCCGCTGGTTTC	9329
3015	LeupheGlyLeuLeuLeuPheValGlyValGlyLeuPheLeuLeuProAlaArg	3033
9330	TGGTTTTGCCCTACTCTCTGCTGCGAGGGTAGGCATCTACTCTCTCCCAACCGA	9386
RESULT 9		
AAC86646		
IID	AAC86646 standard; DNA; 9611 BP.	
XX	AAC86646;	
XX		
XX	02-APR-2001 (first entry)	
XX	Nucleotide sequence of chimeric Hepatitis C virus clone pH77 (p7)-CV-J6S.	
DE	HCV; HCV strain HC-J6CH; HCV genotype 2a; antiviral; vaccine; ss.	
XX	Synthetic.	
OS	Hepatitis C virus.	
XX		
FF	Key	Location/Qualifiers
FT	CDS	342..9389
FT		/*tag= a
XX		
PN	WO200075338-A2.	
XX		
PD	14-DEC-2000.	
XX		
PF	02-JUN-2000; 2000WO-US15446.	
XX		
PR	04-JUN-1999; 99US-0137693.	
XX	(USSH) US DEPT HEALTH & HUMAN SERVICES.	
PA	Yanagi M, Bukh J, Emerson SU, Purcell RH;	
PI		
DR	WPI; 2001-061728/07.	
DR	P-PSDB; AAB30731.	
XX		
PT	Nucleic acid molecule encoding human hepatitis C virus of genotype 2a	
PT	for developing vaccines, for diagnosis of hepatitis C virus and in	
PT	screening assays for identification of antiviral agents -	
XX		
PS	Disclosure; Page 115-118; 167pp; English.	
XX		
CC	AAC86645-48 represent chimeric cDNA clones of infectious Hepatitis C	
CC	virus (HCV). In each clone the C, E1 and E2 genes are derived from HCV	
CC	strain HC-J6CH genotype 2a. The p7 protein was derived from HCV strain	
CC	HC-J6CH or HCV strain pcV-H77C, and the NS genes were derived from HCV	
CC	strain pcV-H77C. Such HCV sequences are capable of expressing the virus	
CC	when transfected into cells. The HCV protein is useful for assaying	
CC	candidate antiviral agents for activity against HCV. Antibodies specific	
CC	for HCV polypeptide are useful in prevention and treatment of diseases	
CC	caused by HCV in animals, in particular humans. The HCV polypeptides	
CC	serve as immunogens in the development of vaccines for preventing HCV	
CC	in mammals or as antigens in diagnostic assays for detecting the	

1941 AlaAlaValGlySerIleGlyLeuGlyLysValLeuValAspIleLeuAlaGlyTyrGly 1860
 5862 GCCCCATCGGACGCTGGAGTGGGAGAGTCTCTGGGACATCTTTCAGGGATGGC 5921
 1861 AlaGlyIleSerGlyAlaLeuValAlaPheLysIleMetSerGlyGluLysProSerMet 1880
 5922 CGCGGCTGCGGGAGCTCTGTAGCATTCATGAGCATGAGCGGTGCGGCTCCACG 5981
 1881 GluAspValValAlaLeuLeuProGlyIleLeuSerProGlyAlaLeuValValGlyVal 1900
 5982 GAGACATCTGCTCAATCTGCTGCCGCCATCTCTGCTCGAGCCCTTGTAGTGGGTG 6041
 1901 IleCysAlaAlaIleLeuArgHisValGlyProGlyGluGlyAlaValGlnTrpMet 1920
 6042 GTCTGCGCAGCAATCTGCGCGGCGGAGTGGCGGCGGCGGCGGCGGCGGCGGCGG 6101
 1921 AsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaProThrHisTyrValThr 1940
 6102 AACCGGCTAATAGCTTTCGCTCCGCGGGAACCATGTTCGCCCGCAGCACTACGTGCGG 6161
 1941 GluSerAspAlaSerGlnArgValThrGlnLeuLeuGlySerIleThrSerLeu 1960
 6162 GAGAGCGATCGAGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6221
 1961 LeuArgArgLeuHisAsnTrpIleThrGluAspCysProIleProCysGlyGlySerTrp 1980
 6222 CTGAGCGCATGTCATGTCAGTGAATAGCTCGAGGTGATCCACTCATGCTCGGCTGCTG 6281
 1981 LeuArgAspValTrpAspTrpValCysThrIleLeuThrAspPheLysAsnTrpLeuThr 2000
 6282 CTAAGGACATCTGGGACTCGGATGATGAGGTGCTGAGCGACTTAAGACCTGGCTGMAA 6341
 2001 SerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTyrLys 2020
 6342 GCCAAGCTCATGCCAAGCTGCTGGGATTCCTTTGTGCTGCGGCGGCGGATAGG 6401
 2021 GlyValTrpAlaGlyThrGlyIleMetThrThrArgCysProCysGlyAlaAsnIleSer 2040
 6402 GGGGTCTGGCGAGGAGCGCATATGTCACACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6461
 2041 GlyAsnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTrp 2060
 6462 GGACATGTCAAAACGGGAGCATGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6521
 2061 GlnGlyThrPheProIleAsnCysTyrThrGluGlyGlnCysValProLysProAlaPro 2080
 6522 AGTGGGACGTTCCCATTAACGCTACACCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 6581
 2081 AsnPheLysValAlaIleTrpArgValAlaAlaSerGluTyrAlaGluValThrGlnHis 2100
 6582 AACTATAAGTTCGGCTGGAGGGTGTCTGACAGAGATACGTGAGATAGGCGGGTGG 6641
 2101 GlySerTyrHisTyrIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeu 2120
 6642 GGGGACTTCCACTACGTATGCTGGTATGACTACTGACAACTTAAATCCCGTSCAGATC 6701
 2121 ProSerProGluPhePheSerTrpValAspGlyValGlnIleHisArgPheAlaProThr 2140
 6702 CCATCGCCCGAATTTTACAGAAATTCGACGGGGTGGCTACACAGGTTTGGCGCCCT 6761
 2141 ProLysProPhePheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal 2160
 6762 TGCAGCCCTTGTGCGGGAGGAGGTATCATTCAGAGTAGGATCCACAGTAGTACCGGG 6821
 2161 GlySerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThr 2180
 6822 GGGTGCAGATTAATCTTGGCGGCGGCGGAGCGGAGCGGTGTTGACGTTCCATGCTCACT 6881
 2181 AspProSerHisIleThrAlaGluThrAlaAlaArgArgLeuAlaArgGlySerProPro 2200
 6882 GATCCCTCCATATAACAGCAGAGGCGGCGGAGGAGGTTGGCGAGGCTCACCCCT 6941

2201 SerGluAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThr 2220
 6942 TCTATGCCAGCTCTCTCGCTAGCAGCTGCTCGCTCCATCTCTCAAGGCACTTGCACC 7001
 2221 ThrHisGlyLysAlaTyrAspValAspMetValAspAlaAsnLeuPhe 2236
 7002 GCCAACCATGACTCCCTCGACCCGAGCTCATAGAGCTAACTCTGTGGAGCAGGAG 7061
 2237 MetGlyGlyAspValThrArgIleGluSerGlySerLysValValValLeuAspSerLeu 2256
 7062 ATGGGCGGCAACATCAGGAGGTGTAGTCAAGAACAAAGTGTGTATCTTGGACTCTTC 7121
 2257 AspProMetValGluGluArgSerAspLeuGluProSerIleProSerGluTyrMetLeu 2276
 7122 GATCCGCTTGTGCGAGGAGGATGAGCGGAGGTCTCGTACCTCAGAGAAATCTCGCG 7181
 2277 ProLysLysArgPheProProAlaLeuProAlaTrpAlaArgProAspTyrAsnProPro 2296
 7182 AAGTCTCGGAGATTCGCCCGGCTCTGCGCTGCGGCGGCGGCGGAGCTACACCCCG 7241
 2297 LeuValGluSerTrpLysArgProAspTyrGlnProAlaThrValAlaGlyCysAlaLeu 2316
 7242 CTAGTAGAGCGTGGAAAGCTGACTACGAAACCACTGTGTCTCATGGCTGCCGCTA 7301
 2317 ProProArgLysThrProThrProProArgArgArgArgThrValGlyLeuSer 2336
 7302 CCACCTCCAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7361
 2337 GluAspSerIleGlyAspAlaLeuGlnLeuAlaIleLysSerPheGlyGlnProPro 2356
 7362 GAATCAACCTATCTACTGCTTGGCGGAGTTCGCCCAAAAGTTTGGCAGCTCTCTCA 7421
 2357 ProSerGlyAspSerCysLeuSerThrGlyAlaGlyAlaAlaAspSerGlySerGlnThr 2376
 7422 ACTTCCGCTATTCGCGGAGTTCGCGGAGTTCGCGGAGTTCGCGGAGTTCGCGGAG 7481
 2377 ProProAspGluLeuAlaLeuSerGluThrGlySerIleSerSerMetProProLeuGlu 2396
 7482 CCCCCGAC-----TCCGAGGTGAGTCTCTATTTCTTCCATGCCCCCTCGGAG 7529
 2397 GlyGluLeuGlyAspProAspLeuGluProGluGlnValGluProGlnProProGln 2416
 7530 GGGGAGCTTGGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7553
 2417 GlyGlyValAlaAlaProGlySerAspSerGlySerThrSerThrCysSerGlu 2434
 7554 -----AGCGAGGCTCATGTCTGAGCGCTAGTAGTGGGGCC 7589
 2435 GluAspAspSerValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuThr 2454
 7590 GACCGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7649
 2455 ProCysSerProGluGluGluLysLeuProIleAsnProLeuSerAsnSerLeuLeuArg 2474
 7650 CGGTGCTGCGGAGAACAAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7709
 2475 TyrHisAsnLysValTyrCysThrThrThrLysSerAlaSerLeuArgAlaLysVal 2494
 7710 CATCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7769
 2495 ThrPheAspArgMetGlnValLeuAspSerTyrTyrAspSerValLeuLysAspIleLys 2514
 7770 AATTTGACAGCTGCAAGTTCGACGAGCAATACGAGGAGTGTCTCAAGGAGGTCAAA 7829
 2515 LeuAlaSerLysValThrAlaArgLeuLeuThrMetGluGluAlaCysGlnLeuThr 2534
 7830 GCAGCGGCTCAAAAGTGAAGCTAACTGCTATCGTAGAGGAGCTTGCAGCTCTGACG 7889
 2535 ProProHisSerAlaArgSerLysTyrGlyPheGlyAlaLysGluValArgSerLeuSer 2554
 7890 CCCCCACATTCAGCAAAATCCAAAGTTTGGCTATGCGGCAAAAGAGCTCGCTGCTGCT 7949
 2555 GlyArgAlaValAsnHisIleLysSerValTrpLysAspLeuLeuAspSerGluThr 2574


```
Db 3642 GTCATGCAGATGTATACCAATGTGGACCAAGACCTTGTGGGCTGGCCGCTCCTCAAGGT 3701
Qy 1121 ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArgAsn 1140
Db 3702 TCCCGCTCATGTGACACCTGTACTCGGGCTCTCTCGGACCTTTACCTGGTGCAGAGCAC 3761
Qy 1141 AlaAspValIleProAlaArgArgArgGlyAspLysArgGlyAlaLeuLeuSerProArg 1160
Db 3762 GCCGATGTCAATCCCGTGGCCGCGGAGGTGATAGCAGGGGTAGCCTGCTTTCCGCCCGG 3821
Qy 1161 ProLeuSerThrLeuLysGlySerSerGlyProValLeuCysProArgGlyHisAla 1180
Db 3822 CCCATTTCCTACTTGAAGGCTCTCGGGGGTCCGCTGTGTGTGCCCGGGACACGCC 3881
Qy 1181 ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle 1200
Db 3882 GTGGGCTATTTCAGGGCCGCGGTGTGCCCGGTGGAGTGGCTAAAGCGGTGGACTTTATC 3941
Qy 1201 ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro 1220
Db 3942 CCTGTGGAGAACCTAGGGACCAACATCAGATCCCGGTGTTCACGGACAACCTCTCTCCA 4001
Qy 1221 ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLys 1240
Db 4002 CCAGCAGTCCCCAGAGCTTCCAGGTGGCCACCTGCATGCTCCACCGGCGAGCGTAAG 4061
Qy 1241 SerThrLysValProValAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnPro 1260
Db 4062 AGCACCAAGGTCCCGCTGCGTACGACGCCAGGGCTACAGAGGTGTGGTGTCAACCCC 4121
Qy 1261 SerValAlaAlaThrLeuGlyPheGlyValAlaTyrLeuSerLysAlaHisGlyIleAsnPro 1280
Db 4122 TCTGTTGCTGCAACGCTGGCTTTGGTGTTCATACATCCCAAGGCCCATGCGGTGATCCT 4181
Qy 1281 AsnIleArgThrGlyValArgThrValThrThrGlyAlaProIleThrTyrSerThrTyr 1300
Db 4182 AATATCAGGACCGGGTGAGAACAAATTAACCATGCGCACCGCCCATCAGCTACTCCACCTAC 4241
Qy 1301 GlyLysPheLeuAlaAspGlyGlyCysAlaGlyValAlaTyrAspIleIleCysAsp 1320
Db 4242 GGCAGGTTCCTTGGCGAGCGGGGTGCTCAGGAGGTGCTTATGACATAATAATTGTGAC 4301
Qy 1321 GluCysHisAlaValAspSerThrIleLeuGlyIleGlyThrValLeuAspGlnAla 1340
Db 4302 GAGTGCACCTCCAGGATGCCACATCCATCTTGGGCATCGGCATGCTCTTGACCAAGCA 4361
Qy 1341 GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProGlySerValThr 1360
Db 4362 GAGACTCGGGGGGAGACTGTTGTGTGCTGCCACTGCTACCCCTCCGGGCTCCGTCACCT 4421
Qy 1361 ThrProHisProAsnIleGluGluValAlaLeuGlyGlnGluGlyGluIleProPheTyr 1380
Db 4422 GTGTCCCATCCTAAACATCGAGGAGGTGTCTCTGTCCACCCGAGAGATCCCCCTTTAC 4481
Qy 1381 GlyArgAlaIleProLeuSerTyrIleLysGlyGlyArgHisLeuIlePheCysHisSer 1400
Db 4482 GGCAGGCTATCCCTCTCGAGTGTATCNAAGGGGGGAGACATCTCATCTTCTGCCACTCA 4541
Qy 1401 LysLysLysCysAspGluLeuAlaAlaLeuArgGlyMetGlyLeuAsnSerValAla 1420
Db 4542 AAGAAGAAGTGCAGCAGACTCGCCGCGAAGTGTGTCGTCATTTGGGCATCAATCGCGTGGCC 4601
Qy 1421 TyrTyrArgGlyLeuAspValSerValIleProThrGlnGlyAspValValValAla 1440
Db 4602 TACTACCGCGGTCTGTAGCTGTGTATCCCGACCGCGCGAGTGTGTGCTGTGTGTCG 4661
Qy 1441 ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnVal 1460
Db 4662 ACCGATGCTCTCATGACTGGCTTTACCGGCGACTTCGACTCTGTGATAGACTGCAACACG 4721
Qy 1461 AlaValThrGlnValAlaAspPheSerLeuAspProThrPheThrIleThrGlnIle 1480
```

QY	381	ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr	400	MetLeuIleLeuLeuGlyGlnAlaGluAlaLeuLeuGlyLeuValIleLeuHisAla	760
DB	1482	GTGGACGGCGCACCATATACTGTTGGGGTTCTCCCGCGCAGACACCGCGCCCTCAC	1541	ATGCTCATCTGTTGGCGCAGCCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	2621
QY	401	SerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTrp	420	AlaSerAlaAlaSerCysAsnGlyPheLeuThrPheValIlePhePheValAlaAlaTrp	780
DB	1542	AGCTTATTTGACATGGCGCCCGCAGGAGAAATCCAGCTCGTTAACACCAATGGCAGCTG	1601	GGAGCGCAGCTAGCTGCATGGCTTCTTATTTTGTCTCATCTTTTGTGGCTGCTGG	2681
QY	421	HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer	440	TyrIleLysGlyArgValValProLeuAlaThrTyrSerLeuThrGlyLeuTrpSerPhe	800
DB	1602	CACATCAACCGCACCGCCCTGAACTGCAATGACTCTCTGCACACCGGCTTATTCGGCTCT	1661	TACATCAAGGGTCGGGTAGTCCCTTAGCTACCTATTCCTCTACCTGGCTGGTCTCTT	2741
QY	441	LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg	460	SerLeuLeuLeuAlaLeuProGlnGlnAlaTyrAlaTyrAspAlaSerValHisGly	820
DB	1662	CTGTTCTACACCCACAGCTTCACTGTCAGGATGTCGCCAGACGATGTCGCGCTGCCG	1721	AGCTTACTGCTCTAGCATTTGCCCAACAGCATATGCACTGGACACGGAGGTGGCGCG	2801
QY	461	SerIleGluAlaPheArgValGlyTrpGlyAlaLeuGlnTyrGluAspAsnValThrAsn	480	GlnIleGlyAlaAlaLeuLeuValMetIleThrLeuPheThrLeuThrProGlyTyrLys	840
DB	1722	AGTATCGAGGCTTCCGGGTGGGATGGGGCGCTTGCATATGAGGATAATGTCACCAAT	1781	TCGTGTGGCGGGTGTCTTCTGTCGGGTTAATGGCGCTGACTCTGTGCGCATATTACAAG	2861
QY	481	ProGluAspMetArgProTyrCysTrpHisTyrProProArgGlnCysGlyValValSer	500	ThrLeuLeuSerArgPheLeuTrpTrpLeuCysTyrLeuLeuThrLeuGlyGluAlaMet	860
DB	1782	CCAGAGGATATGAGACCTTATGCTGCACTACCCACCAAGCAGTGTGGCGTCTCC	1841	CGCTATATCAGCTGGTGCATGGTGGCTTCAGTATTTCTGACCAAGTAGAAGCGCAA	2921
QY	501	AlaLysThrValCysGlyProValTyrCysPheThrProSerProValValGlyThr	520	ValGlnGluTrpAlaProProMetGlnValArgGlyGlyArgAspGlyIleIleTrpAla	880
DB	1842	GCAGAGACTGTGTGGCCAGTGACTGTTTCAACCCAGCCAGTGTGTAGTGGCAGC	1901	CTGCAGCTGGGTTCCCCCTCAACGTCGCGGGGGGGCGGATGCTCATCTTACTTC	2981
QY	521	ThrAspArgLeuGlyAlaProThrTyrThrTrpGlyGluAsnGluThrAspValPheLeu	540	ValAlaIlePheTyrProGlyValValPheAspIleThrLysTrpLeuLeuAlaValLeu	900
DB	1902	ACCGACAGGCTTGGAGCGCCCACTTACACGTGGGGGAGAAATGAGACAGATGCTTCTTA	1961	ATGTGTGTAGTACACCGCACCTGTGTATTTGACATCACCAAACTACTCTCTGGCCATCTC	3041
QY	541	LeuAsnSerThrArgProProLeuGlySerTrpPheGlyCysThrTrpMetAsnSerSer	560	GlyProAlaTyrLeuLeuLysGlyAlaLeuThrArgValProTyrPheValArgAlaHis	920
DB	1962	TTGAACAGCACTCGACACCGCTGGGGTATGCTTGGCTGCACGTGGATGAATCTTCT	2021	GGACCCCTTGGATCTTCAAGCCAGTTTGTAAAGTCCCTTACTCTGCGCGCTTCAA	3101
QY	561	GlyTyrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer	580	AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyGlyArgTyrValGlnMet	940
DB	2022	GGCTACACCAAGACTTGGCGCGCACCACTCGCTGCTAGAGCTGACTTCAACGCCAGC	2081	GGCTTCTCCGATCTGCGCGTAGCGCGGAGATAGCCGAGGTCAATTACGTGCAATG	3161
QY	581	ThrAspLeuLeuCysProThrAspCysPheArgLysHisProAspThrThrTyrLeuLys	600	AlaLeuAlaLeuGlyArgTrpThrGlyThrTyrIleTyrAspHisLeuThrProMet	960
DB	2082	ACGGACCTGTTGGCCCCAGGACTGTTTATAGGAGACCTCTGATACCACTTACCTCAA	2141	GCCATCATCAAGTTAGGGCGCTTACTGGCACCTATGTATTAACCATCTCACCCCTCT	3221
QY	601	CysGlySerClyProTrpLeuThrProArgCysLeuIleAspTyrProTyrArgLeuTrp	620	SerAspTrpAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIleIlePhe	980
DB	2142	TGGCGCTCTGGGCGCTGCTCAGCCCAAGGTGCTGATCGACTACCCCTACAGGCTCTG	2201	CGAGACTGGGGCCCAACAGCGCTGGGAGATCTGGCGGTGGCTGTGGAAACAGTCGTCTC	3281
QY	621	HisTyrProCysThrValAsnTyrThrIlePheLysIleArgMetTyrValGlyGlyVal	640	SerProMetGluLysLysValIleValTrpGlyAlaGluThrAlaAlaCysGlyAspIle	1000
DB	2202	CATTACCCCTGCACAGTTAACTATACATCTTCAAAATAGGATGATGTGGAGGGGT	2261	TCCCGAATGGAGACCAAGCTCATCAGTGGGGGGCAGATACCGCCCGTGGGTGACATC	3341
QY	641	GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp	660	LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp	1020
DB	2262	GAGCACAGGCTCACGGCTGCATCAATTTTCACTCGTGGGATCGTTGCAACTTGGAGGAC	2321	ATCAACGGCTTCCCGCTCTCTCCCGCTAGGGCCAGAGATATCTGTTGGCCAGCCGAC	3401
QY	661	ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGluTrpAlaIleLeuPro	680	GlyTyrThrSerLysGlyTrpSerLeuLeuAlaProIleThrAlaTyrAlaGlnGlnThr	1040
DB	2322	AGAGACAGAGTCAACTGCTCTCTTGTGTGCACTCCACCAAGGAAATGGGCCATTTTACCT	2381	GGAAATGGTCTCAAGGGGTGGAGGTGTGTGGCGCCCATCACGGCGTACGCCAGCAGACG	3461
QY	681	CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisLeuHisGlnAsnIle	700	ArgGlyLeuLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGluGlnAla	1060
DB	2382	TGCTCTTACTCGGACCTGCCCGCTTGTGCACTGGTCTTCTCCACCTCCACCAAAATC	2441	AGAGGCTCTTAGGGTGTATATATCACCAGCTGACTGGCGGGGACAAAAACCAAGTGGAG	3521
QY	701	ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrLysTyrIleValArgTrp	720	GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly	1080
DB	2442	GTGGAGTCAATTCATGTATGSCCTTATCACCTGCCCTCACAAATATACATCGTCCGATG	2501	GGTGGAGTCCAGATCGTGTCACTGTACCCAAACCTTCTTGGCAACGTGCTCAATGGG	3581
QY	721	GluTrpValIleLeuLeuPheLeuLeuAlaAspAlaArgValCysAlaCysLeuTrp	740	ValLeuTrpThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro	1100
DB	2502	GAGTGGGTATATCTTATTTCTGCTTATAGCGGACGCCAGGGTTTTGCGCTGCTTATG	2561	GTATGCTGAGCTGTCTTACACCGGGCGCGAAACAGAGACCATCGCATCCCAAGGGTCTCT	3641
QY	1101	ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTrpProSerProGly	1120		

Db 9330 TGGTTTGGCTACTCTCTGCTGCGGAGGTTAGGATCTACCTCTCCCAACCGA 9386

RESULT 8
AAC86647
ID AAC86647 standard; DNA; 9611 BP.

AC AAC86647;
XX
DT 02-APR-2001 (first entry)
XX
DE Nucleotide sequence of chimeric Hepatitis C virus clone J6S.
XX
KW HCV; HCV strain HC-J6CH; HCV genotype 2a; antiviral; vaccine; ss.

XX Synthetic.
OS Hepatitis C virus.

OS
PH Key Location/Qualifiers
FT 342..9389
PT CDS /*tag= a

XX
XX W0200075338-A2.
XX
XX 14-DEC-2000.
XX
XX 02-JUN-2000; 2000WO-US15446.
XX
XX 04-JUN-1999; 99US-0137693.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Yanagi M, Bukh J, Emerson SU, Purcell RH;
XX
XX WPI; 2001-061728/07.
XX
XX P-PSDB; AAB30732.

PT Nucleic acid molecule encoding human hepatitis C virus of genotype 2a
PT for developing vaccines, for diagnosis of hepatitis C virus and in
PT screening assays for identification of antiviral agents -

XX Disclosure; Page 131-134; 167pp; English.

XX AAC86645-48 represent chimeric cDNA clones of infectious Hepatitis C
XX virus (HCV). In each clone the C, E1 and E2 genes are derived from HCV
XX strain HC-J6CH genotype 2a. The p7 protein was derived from HCV strain
XX HC-J6CH or HCV strain pcV-H77C, and the NS genes were derived from HCV
XX strain pcV-H77C. Such HCV sequences are capable of expressing the virus
XX when transfected into cells. The HCV protein is useful for assaying
XX candidate antiviral agents for activity against HCV. Antibodies specific
XX for HCV polypeptide are useful in prevention and treatment of diseases
XX caused by HCV in animals, in particular humans. The HCV polypeptides
XX serve as immunogens in the development of vaccines for preventing HCV
XX in mammals or as antigens in diagnostic assays for detecting the
XX presence of HCV in biological samples. The HCV polynucleotide is also
XX useful for identifying cell lines capable of supporting the replication
XX of HCV in vitro and to produce attenuated viral strains via passage in
XX vitro or in vivo.

SQ Sequence 9611 BP; 1935 A; 2896 C; 2682 G; 2098 T; 0 other;

Alignment Scores:
Pred. No.: 0 Length: 9611
Score: 13049.00 Matches: 2395
Percent Similarity: 87.40% Conservative: 261
Best Local Similarity: 78.81% Mismatches: 353
Query Match: 80.95% Indels: 30
DB: 22 Gaps: 4

US-09-980-559-2 (1-3033) x AAC86647 (1-9611)

Oy 1 MetSerThrAsnProGlnArgLysThrLysArgAsnThrAsnArgProGln 20
|||||

342 ATGAGCAAAATCTAAACCTCAAAAGAAAAACAAAGAAAAACCAACCGTCGCCACAA 401
Qy
21 AspValLysPheProGlyGlyGlnIleValGlyGlyValTyLeuLeuProArgAsp 40
Db
402 GACGTTAAAGTTTCCGGCGCGCGCAGATCGTTGGCGGAGTATATCTTGTTCGGCGAG 461
Qy
41 GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly 60
Db
462 GCGCCAGGTTGGGTGTGCGCGCGCAAGAAAGACTTCGGAGCGGTCCCGCAGCACGTGA 521
Qy
61 ArgArgGlnProIleProLysAspArgSerThrGlyLysSerTrpGlyLysProGly 80
Db
522 AGCGCGCAGCCCATCCCTAAAGATCGGCGCTCCATCGGCAATCTCTGGGAAACACAGGA 581
Qy
81 TyrProTrpProLeuTyArgLysGluGlyValAlaGlyTyAlaGlyTyTrpLeuLeuSerPro 100
Db
582 TACCCCTGGCCCTATACGGGAATGAGGACTCGGCTGGCAGGATGGCTCTGTCTCCCC 641
Qy
101 ArgGlySerArgProSerTrpGlyProAsnAspProArgHisArgSerArgHisValGly 120
Db
642 CGAGGTTCCCGTCCCTCTTGGGGCCCAATGACCCCGCATAGGTGCGCAACGTGGGT 701
Qy
121 LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyTrpIleProValVal 140
Db
702 AAGGTCATCATACCTAACGTGCGGCTTTGCCGACCTCATGGGGTACATCCCTGTCTGTG 761
Qy
141 GlyAlaProLeuGlyGlyValAlaAlaArgAlaLeuAlaHisGlyValArgValLeuGluAsp 160
Db
762 GCGCGCCCGCTCGCGCGCGCTCGCGAGAGCTCTCGCGCATGGGTGAGAGTCTCTGGAGGAC 821
Qy
161 GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla 180
Db
822 GGGGTTAAATTTGCAACAGGAACTTACCCGGTGGCTCTCTTTCTATCTCTTCTGTGGCC 881
Qy
181 LeuLeuSerCysIleThrThrProValSerAlaAlaGluValLysAsnIleSerThrGly 200
Db
882 CTGCTGCTGTCATCACACCCCGGTCTCGCTGCCAAGTGAAGAACATCATGATCCGGC 941
Qy
201 TyrMetValThrAsnAspCysThrAsnAspSerIleThrTrpGlnLeuGlnAlaVal 220
Db
942 TACATGGTGACTAACGACTGACCAATGACAGATTTACTTGGCAGCTCCAGGCTGCTGTC 1001
Qy
221 LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTrpIle 240
Db
1002 CTCACAGTCCCCGGGTGCGTCCGTCGAGAAAGTGGGAATGCATCTCAGTCTGGATA 1061
Qy
241 ProValSerProAsnValAlaValGlnArgProGlyAlaLeuThrGlnGlyLeuArgThr 260
Db
1062 CCGGTCTCACCGAATGTGCGCGTGCAGCGCGCGCCCTCACGAGGGCTTTGGCGAGC 1121
Qy
261 HisIleAspMetValValMetSerAlaThrLeuCysSerAlaLeuTyValGlyAspLeu 280
Db
1122 CACATGCACATGTTGTGATGTCGCGCAGCTCTGCTGCGCTCTACGTGGGGGACCTC 1181
Qy
281 CysGlyGlyValMetLeuAlaAlaGlnMetPheIleValSerProGlnHisHisTrpPhe 300
Db
1182 TGGCGTGGGTGATGCTCGCAGCCCAATGTTTCAATGTTCTCGCGCGCAGCACCATCTG 1241
Qy
301 ValGlnAspCysAsnCysSerIleTyProGlyThrIleThrGlyHisArgMetAlaTrp 320
Db
1242 GTCCAGAGCTGCAATTTGCTCCATCTACCTCTGGTACCATCACTGGACACCCGATGGC 1301
Qy
321 AspMetMetMetAsnTrpSerProThrAlaThrMetIleLeuAlaTyAlaMetArgVal 340
Db
1302 GACATGATGATGAACATGGTCCCGCAGCGCTACCATGATCTTGGCGTACGGATGCGGTCTC 1361
Qy
341 ProGluValIleIleAspIleIleSerGlyAlaHisTrpGlyValMetPheGlyLeuAla 360
Db
1362 CCGGAGGTCAATATAGACATCATTAGCGGGGCTCATTTGGGGGCTCATGTTGGGTGGCC 1421
Qy
361 TyrPheSerMetGlnGlyAlaTrpAlaLysValValIleLeuLeuLeuAlaAlaGly 380
Db
1422 TACTTCTATGCGAGGAGCGTGGGGAAGTGTGTTGTCTATCTCTTGTGTGGCGCGCGG 1481

2297 LeuValGluSerTrpLysArgProAspTyrGlnProAlaThrValAlaGlyCysAlaLeu 2316
2742 CTAGTAGAGACGTGGAAAAAGCCTGACTACGAACACCTGTGTCTCCATGGCTGCCGCTA 7301
2317 ProProProArgLysThrProThrProProProProProProProProProProProPro 2336
7302 CCACCTCCAGCGTCCCTCTCTGTCCTCCGCTCCGAAAGAGCTAGCGTGGTCTCACC 7361
2337 GluAspSerIleGlyAspAlaLeuGlnGlnLeuAlaIleLysSerPheGlyGlnProPro 2356
7362 GAATCAACCTATCTACTGCTTGGCGAGCTTCCACCAAAAGTTTGGAGCTCCTCA 7421
2357 ProSerGlyAspSerGlyLeuSerThrGlyAlaGlyAlaAlaAspSerGlySerGlnThr 2376
7422 ACTTCCGGCATACGGGCGACATACGACACATCTCTCAGCGCCGCTTCTGGCTGC 7481
2377 ProProAspGluLeuAlaLeuSerGluThrGlySerIleSerSerMetProProLeuGlu 2396
7482 CCCCCCGAC-----TCCGAGCTTGAGTCTCTATTCTTCCATGCCCCCTGGAG 7529
2397 GlyGluLeuGlyAspProAspLeuGluProGluGlnValGluProGlnProProGln 2416
7530 GGGGAGCCTGGGATCCGATCTC----- 7553
2417 GlyGlyValAlaAlaProGlySerAspSerGlySerTrpSerThrCysSerGlu----- 2434
7554 -----ACGCGAGGTGATGTCGACGCTCAGTAGTGGGGC 7589
2435 GluAspAspSerValValCysCysSerMetSerTrpThrGlyAlaLeuIleThr 2454
7590 GACACGGAAGATGCTGTGCTCTCAATGCTTATTCTGGACAGCGCACTCGTCCACC 7649
2455 ProCysSerProGluGluGluLysLeuProIleAsnProLeuSerSerAsnSerLeuLeuArg 2474
7650 CCGTGGCTCGGGAAGAACAAACCTGCCATCAACGCACTGAGCAACTCGTTGCTACGC 7709
2475 TyrHisAsnLysValTyrCysThrThrThrLysSerAlaSerLeuArgAlaLysLysVal 2494
7710 CATCAAACTGTGTATTCACCACTTCAGCAGCTGCTGCCAAGGAGGAGAAAGTC 7769
2495 ThrPheAspArgMetGlnValLeuAspSerTyrTyrAspSerValLeuLysAspIleLys 2514
7770 ACATTTGACAGACTGCAAGTCTTGACAGCATTACAGGACGCTGCTCAAGGAGGTCAA 7829
2515 LeuAlaAlaSerLysValThrAlaArgLeuLeuThrMetGluGluAlaCysGlnLeuThr 2534
7830 GCAGCGCGCTCAAAAGTGAGGCTAATCTGCTATCCGTAGAGGAAGCTTGCAGCCTGACG 7889
2535 ProProHisSerAlaArgSerLysTyrGlyPheGlyAlaLysGluValArgSerLeuSer 2554
7890 CCCCCACATTCACCAATCCAAAGTTTGGCTATGGGGCAAGAGCTGCTGGCCATGCC 7949
2555 GlyArgAlaValAsnHisIleLysSerValTrpLysAspLeuLeuGluAspSerGluThr 2574
7950 AGAAAGCGCTAGCCACATCAACTCCGTTGGTGAAGACCTTCTGGAAGACAGTGAACA 8009
2575 ProIleProThrThrIleMetAlaLysAsnGluValPheCysValAspProThrLysGly 2594
8010 CCAATAGACACTACCACTATGCGCAAGACGAGGTTTCTGGCTTCAGGCTGAGAGGGG 8069
2595 GlyLysLysAlaAlaArgLeuIleValTyrProAspLeuGlyValArgValCysGluLys 2614
8070 GGTCTGTAACCGAGCTCGTCTATCGGTTCCTCCGACCTGGCGTGGCGGTGGCGAAG 8129
2615 MetAlaLeuTyrAspIleThrGlnLysLeuProGlnAlaValMetGlyAlaSerTyrGly 2634
8130 ATGGCCCTGTACGACGTGTAGCAAGCTCCCTCCGCTGATGGGAAGCTCCTACGGA 8189
2635 PheGlnTyrSerProAlaGlnArgValGluPheLeuLeuLeuAlaTrpAlaGluLysLys 2654
8190 TTCCAAATCTACACAGGAGCGGGTTGAATTCCTCTGTCAGCGGTGGAAGTCCAAGAG 8249

2655 AspProMetGlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluArgAsp 2674
8250 ACCCGAGTGGGTTCGTATGATATACCGCTGTTTGTGCTCCACAGTCACTGAGAGCGAC 8309
2675 IleArgThrGluGluSerIleTyrArgAlaCysSerLeuProGluGluAlaHisThrAla 2694
8310 ATCCGTCAGGAGGAGCAATTACCAATGTTGTGACTTGAGCCGCCCAACCCGCGTGGCC 8369
2695 IleHisSerLeuThrGluArgLeuTyrValGlyGlyProMetPheAsnSerLysGlyGln 2714
8370 ATCAAGTCCCTCAGTACGAGGCTTATGTTGGGGCCCTCTTACCAATTCAGAGGGGAA 8429
2715 ThrCysGlyTyrArgCysArgAlaSerGlyValLeuThrThrSerMetGlyAsnThr 2734
8430 AACTCGGCTACCGCAGGTGCGCGAGCGGCTACTGACAACTAGCTGTGTGTACACC 8489
2735 IleThrCysTyrValLysAlaLeuAlaCysLysAlaAlaGlyIleIleAlaProThr 2754
8490 CTCACCTGTACATCAAGCGCGGCGAGCTGTCGAGCGCAGGCTCCAGGACTGCACC 8549
2755 MetLeuValCysGlyAspLeuValIleSerGluSerGlnGlyThrGluGluAsp 2774
8550 ATGCTCGTGTGGCGAGCTTAGTCTGTTATCTGTGAAGTGGCGGGTCCAGGAGGAC 8609
2775 GluArgAsnLeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAsp 2794
8610 GCGCGAGCTGAGAGCTTACGAGGCTATGACAGGTACTTCGCCCCCCCCGGGAC 8669
2795 ProProArgProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerVal 2814
8670 CCCCCACAACCAAGATACGACTTGAGCTTATAACATCATGCTCTCTCCACACGTGTCAGTC 8729
2815 AlaLeuGlyProGlnGlyArgArgTyrTyrLeuThrThrArgAspProThrProIle 2834
8730 GCCCAGCGCGCTGGAAAGAGGCTACTACTTACCCGTCGACCTTACCAACCCCTC 8789
2835 AlaArgAlaAlaTrpGluThrValArgHisSerProValAsnSerTrpLeuGlyAsnIle 2854
8790 GCGAGAGCGCTGGGAGACACAGACACACTCCAGTCAATCTCTGCTAGGCAACATA 8849
2855 IleGlnTyrAlaProThrIleTrpAlaArgMetValLeuMetThrHisPhePheSerIle 2874
8850 ATCATGTTTGGCCCCACACTGTGGCGAGGATGATGATGATGATGATGATGATGATGATGAT 8909
2875 LeuMetAlaGlnAspThrLeuAspGlnAsnLeuAsnPheGluMetTyrClyAlaValTyr 2894
8910 CTCATAGCAGGATGACGCTTGAAACAGGCTCTTAATCTGTGAGATCTACGAGCCTGCTAC 8969
2895 SerValSerProLeuAspLeuProAlaIleIleGluArgLeuHisGlyLeuAspAlaPhe 2914
8970 TCCATAGAACCACTGGATCTACCTCAATCATTCAAAGACTCCATGGCTCTAGCGCATTT 9029
2915 SerLeuHisThrTyrThrProHisGluLeuThrArgValAlaSerAlaLeuArgLysLeu 2934
9030 TCATCTCACAGTACTCTCCAGTCAATCAATAGGTTGGCGCATGCTCTCAGAAAATCT 9089
2935 GlyAlaProProLeuArgAlaTrpLysSerArgAlaArgAlaValArgAlaSerLeuIle 2954
9090 GGGGTCCCGCTTCGAGCTTGGAGCACCGGGCCCGGAGCGTCCGCTAGGCTTCTG 9149
2955 SerArgGlyGlyArgAlaAlaValCysGlyArgTyrLeuPheAsnTrpAlaValLysThr 2974
9150 TCCAGAGGAGCGAGGCTGCTATATGTCGCAAGTACTCTTCACTGGCAGTAGAACA 9209
2975 LysLeuLysLeuThrProLeuProGluAlaArgLeuLeuAspLeuSerSerTrpPheThr 2994
9210 AAGCTCAAACTCACTCAATAGCGCGCTGGCGCGCTGGACTTGTCCGCTGGTTTACG 9269
2995 ValGlyAlaGlyGlyAspIleTyrHisSerValSerArgAlaArgProArgLeuLeu 3014
9270 GCTGCTTACAGCGGGGAGACATTTATCACAGCGTGTCTCATGCCCCCGCTGGTTC 9329
3015 LeuPheGlyLeuLeuLeuPheValGlyValGlyLeuPheLeuLeuProAlaArg 3033

Db	5022	TTTTGGGAGGGCGTCTTTACGGCGCTCACTCATATAGATGCCACATTTTATCCAGACA	5081	6102	AACCGGCTAATAGCCTCCGCTCCCGGGGAAACCAATGTTTCCCCACGCACTACGTGCGG	6161
Qy	1581	LyGlnSerGlyGluAenPheAlaTyrLeuThrAlaTyrGlnAlaThrValCysAlaArg	1600	1941	GluSerAspAlaSerGlnArgValThrGlnLeuLeuGlySerLeuThrIleThrSerLeu	1960
Db	5082	AAGCAGAGTGGGAGAACTTTCTTACCTCGTGTAGCGTACCAAGCCACCGGTGCGGTAGG	5141	6162	GAGAGGATGACCGCGCGCTCACTGCCATACCTCAGCAGGCTCACTGTAAACCCAGCTC	6221
Qy	1601	AlaLysAlaProProSerTrpAspValMetTrpLysCysLeuThrArgLeuLysPro	1620	1961	LeuArgArgLeuHisAsnTrpIleThrGluAspCysProIleProCysGlyGlySerTrp	1980
Db	5142	GCTCAAGCCCTCCCCCATCGTGGACCATGTGGAGTGTGTGATCCGCTTAAACCC	5201	6222	CTGAGCGCACTCATCAGTGGATAAGCTCGGAGTGTACCACTCCATGCTCCGCTTCCTGG	6281
Qy	1621	ThrLeuValGlyProThrProLeuLeuTyrArgLeuGlySerValThrAenGluValThr	1640	1981	LeuArgAspValTrpAspTrpValCysThrIleLeuThrAspPheLysAsnTrpLeuThr	2000
Db	5202	ACCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTCAGATGAATGATCACC	5261	6282	CTAAGGACATCTGGAGCTGGATATGCGAGGTGCTGAGCGACTTTAAGACCTGGCTGAAA	6341
Qy	1641	LeuThrHisProValThrLysTyrIleAlaThrCysMetClnAlaAspLeuGluValMet	1660	2001	SerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTyrLys	2020
Db	5262	CTGAGCACCACCAATCACCAATCATCATGACATGCAATGCTGCGCGGACCTGGAGGTGCTG	5321	6342	GCCAGCTCATGCGCACACTGCTGGGATTCCTTTGTCTGCTGCCAGCGGGTATAGG	6401
Qy	1661	ThrSerThrTrpValLeuAlaGlyGlyValLeuAlaAlaValAlaTyrCysLeuAla	1680	2021	GlyValTrpAlaGlyThrGlyIleMetThrThrArgCysProCysGlyAlaAsnIleSer	2040
Db	5322	ACGAGCACTGGGTGCTCGTTGGCGGCTCTGGCTGCTCTGGCGGGTATTGCTGTCTCA	5381	6402	GGGTCTGGCGGAGAGACGCAATTATGCACACTCGCTGCCACTGTGGAGCTGAGATCACT	6461
Qy	1681	ThrGlyCysValCysIleIleGlyArgLeuHisIleAenGlnArgAlaValAlaPro	1700	2041	GlyAenValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTrp	2060
Db	5382	ACAGCTGCGTGTATAGTGGGAGGATCGTCTGTGCGGAAGCGCGCAATTATACCT	5441	6462	GGACATGTCAAAACCGGAGCATGAGATCGTGGTCTCTAGGACCTGCAGGAACATGGG	6521
Qy	1701	AspLysGluValLeuTyrGluAlaPheAspGluMetGluGluCysAlaSerArgAlaAla	1720	2061	GlnGlyThrPheProIleAenCysTyrThrGluGlyGlnCysValProLysProAlaPro	2080
Db	5442	GACAGGAGGTCTCTACAGGAGTTCGATGAGATGGAAGATGCTCTCAGCACTTACCG	5501	6522	AGTGGAGCGTTCCCATTAACGCTACACACCGGCGCTGCTACTCCCTCTCTGCGCGG	6581
Qy	1721	LeuIleGluGluGlyGlnArgIleAlaGluMetLeuLysSerLysIleGlnGlyLeuLeu	1740	2081	AsnPheLysValAlaIleTrpArgValAlaAlaSerGlyTyrAlaGluValThrGlnHis	2100
Db	5502	TACATCGAGCAGGAGATGCTCGCTGAGCAGTTCAGCAGAGGCGCTCGGCTCTCTG	5561	6582	AACTATAAGTTCGCGTGTGGAGGGTGTCTGCAGAGGAATAGCTGGAGATAAGCGGGTG	6641
Qy	1741	GlnGlnAlaSerLysGlnAlaGlnAspIleGlnProThrValGlnAlaSerTrpProLys	1760	2101	GlySerTyrHisTyrIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeu	2120
Db	5562	CAGACCGGTCGCCCATGACAGAGGTATACCCCTGCTGCTCCAGACCAACTGGCAGAAA	5621	6642	GGGAGCTTCCACTACGTATCGGTATGACTACTGACAACTTAAATGCCGCTGCCAGATC	6701
Qy	1761	ValGluGlnPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeuAla	1780	2121	ProSerProGluPhePheSerTrpValAspGlyValGlnIleHisArgPheAlaProThr	2140
Db	5622	CTCAGAGTCTTTTGGGCGAAGCACATGTGGAAATTTTCATGATGGGATACAACTTGGCG	5681	6702	CCATCGCCCGAATTTTACAGAAATTTGGAGCGGGTGGCGCTACACAGGTTTGGCGCCCT	6761
Qy	1781	GlyLeuSerThrLeuProGlyAsnProAlaValAlaSerMetMetAlaPheSerAlaAla	1800	2141	ProLysProPhePheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal	2160
Db	5682	GGCTGTCAAGCTGCTCGTGGTAAACCCGCCATGCTCTCATTTGATGGCTTTACAGCTGCC	5741	6762	TGCAAGCCCTTGTGCGGAGGAGGTATCATTCAGAGTAGGACTCCACAGTAGTACCGGTG	6821
Qy	1801	LeuThrSerProLeuSerThrThrThrIleLeuLeuAsnIleLeuGlyGlyTrpLeu	1820	2161	GlySerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThr	2180
Db	5742	GTCACGAGCCCACTAACCACTGGCCAAACCCCTCTCTTCAACATATTGGGGGGTGGTG	5801	6822	GGGTCCGAATTACTTTCGAGCGCCGAAACCGGAGCTAGCGGTGTGACGTCCATGCTCACT	6881
Qy	1821	AlaSerGlnIleAlaProProAlaGlyAlaThrGlyPheValValSerGlyLeuValGly	1840	2181	AspProSerHisIleThrAlaGluThrAlaAlaArgArgLeuAlaArgGlySerProPro	2200
Db	5802	GCTGCCAGCTCGCGCGCGCGCGCTACTGCTACTGCTTCTTCAACATATTGGGGGGTGGTG	5861	6882	GATCCCTCCCATATACAGCAGGCGCGCGGAGAGGTTGGCGAGAGGTCACCCCT	6941
Qy	1841	AlaAlaValGlySerIleGlyLeuGlyValLeuValAspIleLeuAlaGlyTyrGly	1860	2201	SerGluAlaSerSerAlaSerGlnLeuSerAlaProSerLeuAlaThrCysThr	2220
Db	5862	CGCCCATCGGAGCGTGTGGAGTGGGAGAGTCTCTGCTGGACATCTTTCAGAGGTATGGC	5921	6942	TCTATGCCAGCTCTCTCGGTAGCGTGTGCTCCATCTCTCAAGGCACTTGCACC	7001
Qy	1861	AlaGlyIleSerGlyAlaLeuValAlaPheLysIleMetSerGlyGluLysProSerMet	1880	2221	ThrHisGlyLysAlaTyrAspValAspMetValAspAlaAsnLeuPhe	2236
Db	5922	CGGGCGTGGCGGAGCTCTTGTAGCATTTCAAGATCATGAGCGGTGAGGTCCCTCCACG	5981	7002	GCCAAACCATGACTCCCTGACCGCGAGCTCATAGGCTTAACCTCTCTGGAGGCGAGGAG	7061
Qy	1881	GluAspValValAsnLeuLeuProGlyIleLeuSerProGlyAlaLeuValValGlyVal	1900	2237	MetGlyGlyAspValThrArgIleGluSerGlySerLysValValValLeuAspSerLeu	2256
Db	5982	GAGGACCTGGTCAATCTGCTGCCCGCCATCTCTCGCTCGGAGCGCTTGTAGTGGGTGTG	6041	7062	ATGGCGGCAACATCACCAGGTTGAGTCAGAGAACAAAGTGTGTGTTCTGGACTCTTC	7121
Qy	1901	IleCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrpMet	1920	2257	AspProMetValGluGluArgSerAspLeuLeuProSerIleProSerGlyTyrMetLeu	2276
Db	6042	GTCTCGGAGCAATACTCGCGCGGACGTTGCGCGGCGAGGGCGAGTGCATGATGATG	6101	7122	GATCCGCTTGTGGCAGAGGAGGATGAGCGGAGGTCTCCGTACCTCGCAAAATTTCTGCGG	7181
Qy	1921	AsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaProThrHisTyrValThr	1940	2277	ProLysLysArgPheProAlaLeuProAlaTrpAlaArgProAspTyrAsnProPro	2296
Db				7182	AAGTCTCGGAGATTTCGCGCGGCTCTGCGCGTGTGGCGCGCGGCGGACTACAACCCCGG	7241

```
QY 841 ThrLeuLeuSerArgPheLeuTrpTrpLeuCysTyrLeuLeuThrLeuGlyGluAlaMet 860
Db 2862 CGCTATATACAGTGGTGCATGTGGTGGCTTCAGTATTTCTGACACGAGTAGAAGCGCAA 2921
QY 861 ValGlnGluTrpAlaProMetGlnValArgGlyGlyArgAspGlyIleLeuTrpAla 880
Db 2922 CCGCAGCTGTGGGTTCGCCCTCAACGTCGGGGGGGGCGGATGCGGTCACTTACTC 2981
QY 881 ValAlaIlePheTyrProGlyValValPheAspIleThrIleTyrTrpLeuAlaValLeu 900
Db 2982 ATGTGTGTATACACCGACCTGGTATTTGACATCAACCAATCTCTCGGCGCATCTTC 3041
QY 901 GlyProAlaTyrLeuLeuLysGlyAlaLeuThrArgValProTyrPheValArgAlaHis 920
Db 3042 GGACCCCTTTGGATCTTCAAGCCAGTTGCTTAAAGTCCCTACTTCGTGGCGGTTCAA 3101
QY 921 AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyGlyArgTyrValGlnMet 940
Db 3102 GGCCTTCTCGGATCTCGCGCTAGCGCGGAAGATAGCCGAGGTCAATTACGTGCAATG 3161
QY 941 AlaLeuLeuAlaLeuGlyArgTyrThrGlyThrTyrIleTyrAspHisLeuThrProMet 960
Db 3162 GCCATCATCAAGTTAGGGCGCTTACTGCGCACCTATGTGTATAACCATCTCAACCCCTCT 3221
QY 961 SerAspTrpAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIleIlePhe 980
Db 3222 CGAGACTGGGGCAGCAACGCTCGGAGATCTGGCGGTGGCTGTGGNACAGTCGTCTTC 3281
QY 981 SerProMetGlnLysLysValIleValTrpGlyAlaGluThrAlaAlaCysGlyAspIle 1000
Db 3282 TCCCGAATGAGACCAAGCTCATCACGTGGGGGGCAGATACCCGCGCGTGGCGTGACATC 3341
QY 1001 LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp 1020
Db 3342 ATCAACGGCTGTGCCGCTCTGCGCGTAGGGCGCAGGAGATCTGCTTGGCGCGAGCGAC 3401
QY 1021 GlyTyrThrSerLysGlyTyrSerLeuLeuAlaProIleThrAlaTyrAlaGlnGlnThr 1040
Db 3402 GGAATGGTCTCCAAAGGGTGGAGGTGTGTGGCCGCCCATCAAGCGGTACGCCAGCAGACG 3461
QY 1041 ArgGlyLeuLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGluGlnAla 1060
Db 3462 AGAGGCTCTCAGGTGTATAATCACCCAGCTGACTGGCGGGGACAAAAACCAAGTGGAG 3521
QY 1061 GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly 1080
Db 3522 GGTGAGGTCCAGATCGGTCACTGCTACCCAAACCTTCTGCGCAACGTGCATCAATGGG 3581
QY 1081 ValLeuTrpThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro 1100
Db 3582 GTATGTGGACTGTCTACCAAGGGCGGGCGGNACGAGGACCATCGCATCAACCAAGGTCTCT 3641
QY 1101 ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTrpProSerProGly 1120
Db 3642 GTCATCGATGTATACCAATGTGGACCAAGACCTTGTGGGCTGGCGGCTCCTCAAGT 3701
QY 1121 ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArgAsn 1140
Db 3702 TCCCGTCTCATACACCTGTACTCTCGGCTCTCGGACCTTTACCTGTGTACAGGCGAC 3761
QY 1141 AlaAspValIleProAlaArgArgGlyAspLysArgGlyAlaValLeuLeuSerProArg 1160
Db 3762 GCCGATGTCTATCCGCGCGCGGAGGTGATAGCAGGGGTAGCTGTCTTCGCCCGG 3821
QY 1161 ProLeuSerThrLeuLysGlySerSerGlyProValLeuCysProArgGlyHisAla 1180
Db 3822 CCCATTCTCTACTTGAAGGCTCTCTCGGGGGTCCGCTGTGTGGTGGCCCGGGGACAGCC 3881
QY 1181 ValGlyValPheArgAlaValCysSerArgGlyValAlaLysSerIleAspPheIle 1200
Db 3882 GTGGGCTTATTAGCGCGCGGTGTGACCCGCGGAGTGGCTAAAGCGGTGGACTTTATC 3941
QY 1201 ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro 1220
Db 3942 CCTGTGGAGAACCTTAGGACCAACCATGAGATCCCCGGTGTTCACGGACAACCTCTCTCCA 4001
QY 1221 ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLys 1240
Db 4002 CCAGCAGTCCCCCAGAGCTTCCAGGTGGGCCACCTGCATGCTCCACCGCAGCGGTAAAG 4061
QY 1241 SerThrLysValProValAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnPro 1260
Db 4062 AGCACCAAGGTCGGGCTCGGTACGAGCCCGGCTTACAGGTGTGGTGTCTCAACCCC 4121
QY 1261 SerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerLysAlaHisGlyIleAsnPro 1280
Db 4122 TCTGTGTCTGCAACGCTGGCTTGGTGTGTATCATGTCCAGGCCCATGGGTGTATCTCT 4181
QY 1281 AsnIleArgThrGlyValArgThrValThrGlyAlaProIleThrTyrSerThrTyr 1300
Db 4182 AATATCAGGACCGGGGTGAGAACAAATTACCACTTGGCAGGCCCATCATCTACCTACCTAC 4241
QY 1301 GlyLysPheLeuAlaAspGlyGlyCysAlaGlyGlyAlaTyrAspIleIleCysAsp 1320
Db 4242 GGCAGGTTCTTCCGCGACGCGGGTGTCTCAGGAGGTGCTTATGACATAATAATTTGTGAC 4301
QY 1321 GluCysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGlnAla 1340
Db 4302 GAGTGGCACTCCACCGATGCCACATCTTGGGCATCGGCACCTGTCTTGCACCAAGCA 4361
QY 1341 GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProGlySerValThr 1360
Db 4362 GAGACTGCGGGGGCGAGACTGGTGTGTGCTCGGCACCTGTCTACCCCTCCGGCTCCGCTCACT 4421
QY 1361 ThrProHisProAsnIleGluGluValAlaLeuGlyGlnGluGlyLysLeuProPheTyr 1380
Db 4422 GTGTCCCATCTTAAACATCAGAGAGGTGTCTGTCCACCCAGGAGAGATCCCTTTTAC 4481
QY 1381 GlyArgAlaIleProLeuSerTyrIleLysGlyGlyArgHisLeuIlePheCysHisSer 1400
Db 4482 GGCAGGCTTATCCCTCTCAGGTGATCAAGGGGGGAGAGACATCTCATCTTCTGCCACTCA 4541
QY 1401 LysLysLysCysAspGluLeuAlaAlaLeuArgGlyMetGlyLeuAsnSerValAla 1420
Db 4542 AAGAAGAAGTGGCAGCAGCTCGCGCAAGCTGGTGCATTTGGGCATCAATCGCGTGGCC 4601
QY 1421 TyrTyrArgGlyLeuAspValSerValIleProThrGlnGlyAspValValValAla 1440
Db 4602 TACTACCGGGTCTTGACGTGTGTCTATCCGACCGAGCGGCGATGTGTGTCGTCGTGTC 4661
QY 1441 ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnVal 1460
Db 4662 ACCGATGCTCTCATGCTGGCTTACCGCGGACTTCGACTCTGTGTATAGACTGCAACACG 4721
QY 1461 AlaValThrGlnValValAspPheSerLeuAspProThrPheThrIleThrGlnIle 1480
Db 4722 TGTGTCACTCAGACAGTCGATTTTCAAGCTTGCCTTACCTTTACCATTCAGACAACACG 4781
QY 1481 ValProGlnAspAlaValSerArgSerGlnArgArgGlyArgThrGlyArgGlyArgLeu 1500
Db 4782 CTCCTCCAGGATGTCTTCCAGGACTCAACCGGGGGCAGGACTGGCGGGGGAGGCCA 4841
QY 1501 GlyIleTyrArgTyrValSerThrGlyGluArgAlaSerGlyMetPheAspSerValVal 1520
Db 4842 GGCATCTATAGATTTGTGGCCCGGGGAGCGCCCTCCGGCATGTTCGACTCTGTCGCTC 4901
QY 1521 LeuCysGluCysTyrAspAlaGlyAlaAlaIleTyrGluLeuThrProSerGluThrThr 1540
Db 4902 CTCTGTAGTGTATGACGCGGGCTGTGTGTGTATAGTCTCAGCGCCCGCAGACTACA 4961
QY 1541 ValArgLeuArgAlaTyrPheAsnThrProGlyLeuProValCysGlnAspHisLeuGlu 1560
Db 4962 GTTAGGCTACGAGCGGTACATGAACACCCCGGGGCTTCCCGTGTGCGAGACCATCTTGA 5021
QY 1561 PheTrpGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThr 1580
```


Db	642	CGAGGTTCCCGTCCCTTGGGGCCCAATGATGCCCCGGCATAGTTCGCGAACGTGGGT	701	Db	1722	AGTATCGAGCGCTTCCTGGGTGGGATGGGGCGCGCTTGTGCAATATGAGGATAATGTCCACCAAT	1781
Qy	121	LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProValVal	140	Qy	481	ProGluAspMetArgProTyrCysTyrHisTyrProArgGlnCysGlyValValSer	500
Db	702	AAGGTCAATGATACCTTACAGTGGGCTTGGCGGCTTCATGGGGTACATCCCTGTCGTG	761	Db	1782	CCAGAGGATATGAGACCTTATTTGCTGGCACTTACCCACCAAGGAGTGTGGCGTGGTCTCC	1841
Qy	141	GlyValProLeuGlyGlyValAlaArgAlaLeuAlaHisGlyValArgValLeuGluAsp	160	Qy	501	AlaLysThrValCysGlyProValTyrCysPheThrProSerProValValGlyThr	520
Db	762	GGCGCCCGCTCGCGGGTGGCGAGCTCTCGCGCATGGCGTGAGAGTCTCTGGAGGAC	821	Db	1842	CGGAAGACTGTGTGGCCAGTACTGTTCACCCCGCCAGCCAGTGTGTAGTGGGACG	1901
Qy	161	GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla	180	Qy	521	ThrAspArgLeuGlyAlaProThrTyrThrTyrGlyGluAsnGluThrAspValPheLeu	540
Db	822	GGGGTTAAATTTGCAACAGGGAATTTACCGGTGTCTCTTTCTATCTTCTGTCTGGCC	881	Db	1902	ACCGACAGGCTTGGAGCGCCCACTTACAGTGGGGGAGAAATGAGACAGATGTCTTCTTA	1961
Qy	181	LeuLeuSerCysIleThrThrProValSerAlaAlaGluValIleAsnIleSerThrGly	200	Qy	541	LeuAsnSerThrArgProProLeuGlySerTyrPheGlyCysThrTyrMetAsnSerSer	560
Db	882	CTGCTGTCTGTCATCACACCCCGGTCTCCGCTGCGGAAGTGAAGAACATCAGTACCGGC	941	Db	1962	TTGAACAGCACTCGACACCGCTGGGTCACTGTTCCGGCTGCACGCTGGATGAACCTCTCT	2021
Qy	201	TyrMetValThrAsnAspCysThrAsnAspSerIleThrTyrGlnLeuGlnAlaVal	220	Qy	561	GlyTyrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer	580
Db	942	TACATGTGACTAACGACTGCACATGACAGCATTTACCTGGCAGCTCCAGGCTGTCTC	1001	Db	2022	GGCTACACCAAGACTTGGCGCGCACCCCTCTCCGTACTAGAGCTGACTTCAACGCCAGC	2081
Qy	221	LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysIlePhe	240	Qy	581	ThrAspLeuLeuCysProThrAspCysPheArgLysHisProAspThrThrTyrLysLeu	600
Db	1002	CTCCAGCTCCCGGGTGGTCCGTCGAGAAAGTGGGAATGCATCTCAGTCTGGATA	1061	Db	2082	ACGGACCTGTGTGCCCGCACGACTGTTTATAGAAAGCATCTCGATACCACTTACCTCAA	2141
Qy	241	ProValSerProAsnValAlaValAlnArgProGlyAlaLeuThrGlnGlyLeuArgThr	260	Qy	601	CysGlySerGlyProTyrLeuThrProArgCysLeuIleAspTyrProTyrArgLeuTyr	620
Db	1062	CCGGTCTCACCGAATGTGGCGTGCAGCGCGCGCGCCCTCACGAGGGCTTGGCGAGC	1121	Db	2142	TGCGGCTCTGGGCCCTGGGCTCAGCCNAGTGCCTGATCGACTACCCCTACAGGCTCTGG	2201
Qy	261	HisLeaspMetValValMetSerAlaThrLeuCysSerAlaLeuTyrValGlyAspLeu	280	Qy	621	HisTyrProCysThrValAsnTyrThrIlePheLysIleArgMetTyrValGlyGlyVal	640
Db	1122	CACATCGACATGGTGTGGTGTCCGCCACCGCTCTGCTCTCCCTCTACGTGGGGGACCTC	1181	Db	2202	CATTACCCCTCGCACAGTAACTATACCATCTTCAAATAAGGATGTATGTGGGAGGGT	2261
Qy	281	CysGlyGlyValMetLeuAlaAlaGlnMetPheIleValSerProGlnHisIleStrPhe	300	Qy	641	GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp	660
Db	1182	TGCGGTGGGTGATGCTCGCAGCCCAATGTTTCACTTCTCGCGCAGCACCACTGGT	1241	Db	2262	GAGCACAGCTCAGCGCTCATGCAATTTCACTCGTGGGGATGCTGTGCAACTTGGAGGAC	2321
Qy	301	ValGlnAspCysAsnCysSerIleTyrProGlyThrIleThrGlyHisArgMetAlaTyr	320	Qy	661	ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGluTyrAlaIleLeuPro	680
Db	1242	GTCCAGACTGCAATTCCTCATCTACCTCGTACCATCATCTGACACCGCATGGCATGG	1301	Db	2322	AGNACAGAGTCACTGCTCTCTTTGTCACCTCCACCAAGGATGGGCCATTTTACCT	2381
Qy	321	AspMetMetAsnTyrSerProThrAlaThrMetIleLeuAlaTyrAlaMetArgVal	340	Qy	681	CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisLeuHisGlnAsnIle	700
Db	1302	GACATGATGAACTGGTCCGCCACCGCTACCATGATCTTGGCTACGCGATGGCTGTC	1361	Db	2382	TGCTCTTACTCGACCTCGCGCTTGTGCTGCTTCTTCCACCTCCACCAAAACATC	2441
Qy	341	ProGluValIleIleAspIleIleSerGlyAlaHisTyrGlyValMetPheGlyLeuAla	360	Qy	701	ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrIleValArgTyr	720
Db	1362	CCCGAGGTCAATATAGACATCAATAGCGGGCTCATTTGGGGCTCATGTTCCGGCTTGGCC	1421	Db	2442	GTGACGTCAATTTCAATGATGGCTTATCACCTGCGCTTACAAAATACATCGTCCGATGG	2501
Qy	361	TyrPheSerMetGlnGlyAlaTyrAlaLysValValIleLeuLeuAlaAlaGly	380	Qy	721	GluTyrValIleLeuLeuPheLeuLeuAlaAspAlaArgValCysAlaCysLeuTyr	740
Db	1422	TACTTCTATGACGGGAGCGTGGGGAAGTCTGTGTCTCTCTCTTGTGGCGCGCGG	1481	Db	2502	GAGTGGGTAAATCTCTTATCTTCTTGTAGCGGAGCGCAGGGTTTGGCGTCTTATGG	2561
Qy	381	ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr	400	Qy	741	MetLeuIleLeuLeuGlyGlnAlaGluAlaLeuGluLysLeuValIleLeuHisAla	760
Db	1482	GTGACCGCGGCACTTATGTTGGGGTGTCTGCCCGGAGACCAACCGCGCTTATCACC	1541	Db	2562	ATGCTCATCTTGTGGCGCAGCGCGAAGCAGCACTAGAGAAGCTGTGCTATCTTGCAGCT	2621
Qy	401	SerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTyr	420	Qy	761	AlaSerAlaAlaSerCysAsnGlyPheLeuTyrPheValIlePhePheValAlaAlaTyr	780
Db	1542	AGCTTATTTGACATGGGCGCCAGCAGAAATCCAGCTGTGTAACCAATGGCAGCTGG	1601	Db	2622	GCAGGCGCAGTACGTGCAATGGCTTCTATATTTTGTCTATCTTTTTCGGGCTGTCTGG	2681
Qy	421	HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer	440	Qy	781	TyrIleLysGlyArgValValProLeuAlaThrTyrSerLeuThrGlyLeuTyrSerPhe	800
Db	1602	CACATCAACCGCACCGGCTTCACTGCAATGACTCTTGTGACACCGGCTTATCGCTCT	1661	Db	2682	TACATCAAGGCTCGGGTAGTCCCTTACCTTACCTATTTCCCTCACTGGCCCTGTGCTT	2741
Qy	441	LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg	460	Qy	801	SerLeuLeuLeuLeuAlaLeuProGlnGlnAlaTyrAlaTyrAspAlaSerValHisGly	820
Db	1662	CTGTTCTACACCCAGCTTCACTCGTACGATGTCGGAACGCATGTCCGCCCTGCGGC	1721	Db	2742	AGCCTACTGCTCTAGCATTTGCCCAACAGGCATATGCACTGGACACGAGGTGGCGCGC	2801
Qy	461	SerIleGluAlaPheArgValGlyTyrGlyAlaLeuGlnTyrGluAspAsnValThrAsn	480	Qy	821	GlnIleGlyAlaAlaLeuLeuMetIleThrLeuPheThrLeuThrProGlyTyrLys	840
Db				Db	2802	TCGTTGGCGCGCTTGTCTTGTGGGTAAATGGCGCTGACTCTGTGCGCATATTACAAG	2861

Db 8562 GCCCTTGCAGCGTGAAGGCTGCRGGATCGTGACCCCTGTTAATGTTGTGTGTGAGAC 8621
 QY 2761 AspleuValValIleSerGluSerGlnGlyThrGluGluAspGlnArgAsnLeuArgAla 2780
 Db 8622 GACCTGGTCTCATCTCAGAGACCAAGGTAAACGAGGAGGACGAGCAACCTGAGAGCT 8681
 QY 2781 PheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProArgProGluTyr 2800
 Db 8682 TTCACGAGGCTATGACCAAGTATTCGGCCCTCCCGGTGACCTTCCAGACCGGAATAT 8741
 QY 2801 AspleuGluLeuLeuThrSerCysSerSerAsnValSerValAlaLeuGlyProGlnGly 2820
 Db 8742 GACTTGAGCTTATAACATCTCTCTCAACAGTATCGGTAGCGTGGACTCTCGGGT 8801
 QY 2821 ArgArgArgTyrTyrLeuThrArgAspProThrProIleAlaArgAlaAlaTyrGlu 2840
 Db 8802 CGCGCGGTGACTTCTTACACAGAGACCTTACACCTTCCATCATCCGAGCTGTGGAA 8861
 QY 2841 ThrValArgHisSerProValAsnSerTyrLeuGlyAsnIleIleGlnTyrAlaProThr 2860
 Db 8862 ACAGTAAGACACTCCCTGTCAATCTTGCTGGCAACATCATCCAGTACGCCCCACA 8921
 QY 2861 IleTyrAlaArgMetValLeuMetThrHisPheSerIleLeuMetAlaGlnAspThr 2880
 Db 8922 ATCTGGGTCCGATGGTGTATTAATGACTCACTTCTTCCATATATTGGCCAGGACACT 8981
 QY 2881 LeuAspGlnAsnLeuAsnPheGluMetTyrGlyAlaValTyrSerValSerProLeuAsp 2900
 Db 8982 CTGAACCAAAATCTCAATTTTGAGATGTACGGGGCAGTATACCTCGGTCAATCCATTAGAC 9041
 QY 2901 LeuProAlaIleLeuGluArgLeuHisGlyLeuAspAlaPheSerLeuHisThrTyrThr 2920
 Db 9042 CTACCGGCCATAATTGAAGGCTACATGGCTTGAAGCCTTTTCACTGCACACATACTCT 9101
 QY 2921 ProHisGluLeuThrArgValAlaSerAlaLeuArgGlyLeuGlyAlaProProLeuArg 2940
 Db 9102 CCCACGAACTCTCAGGGGTGGCAGCACTCTCAGAAACTTTGAGCGCTCCCTCTAGA 9161
 QY 2941 AlaTyrLeuSerArgAlaArgAlaValArgAlaSerLeuLeuSerArgGlyGlyArgAla 2960
 Db 9162 GCGTGAAGAGTCCGGGGGTGCGGTGAGAGCTTCACTATCGCCCAAGAGCGAGGGCG 9221
 QY 2961 AlaValCysGlyArgTyrLeuPheAsnTyrAlaValLeuThrLysLeuLysLeuThrPro 2980
 Db 9222 GCCATTGTGGCGGTACCTCTTCAACTGGCGGTGAAACAAAGCTCAAACTCACTCCA 9281
 QY 2981 LeuProGluAlaArgLeuAspLeuSerSerTyrPheThrValGlyAlaGlyGlyGly 3000
 Db 9282 TTGCCCGAGCGAGCGCCCTGGATTTATCCGGGTGGTTTACCGTGGGGCGCGCGGGGC 9341
 QY 3001 AspIleTyrHisSerValSerArgAlaArgProArgProArgLeuLeuPheGlyLeuLeu 3020
 Db 9342 GACATTATACAGCGGTGCGATGTCGACCGCCGCTATTACTCTTTGCTTACTCTCTA 9401
 QY 3021 LeuPheValGlyValGlyLeuPheLeuLeuProAlaArg 3033
 Db 9402 CTTAGCGTAGGAGTAGGCACTCTTTTACTCTCCCGCTCGG 9440

RESULT 7

ID AAC86645 standard; DNA; 9611 BP.
 XX
 AC AAC86645;
 XX
 DT 02-APR-2001 (first entry)
 XX Nucleotide sequence of chimeric Hepatitis C virus clone pH77CV-J6S.
 DE HCV; HCV strain HC-J6CH; HCV genotype 2a; antiviral; vaccine; ss.
 KW Synthetic.
 XX Hepatitis C virus.
 OS
 XX

FH Key Location/Qualifiers
 CDS 342..9389
 /*tag= a
 XX WO200075338-A2.
 PD 14-DEC-2000.
 XX 02-JUN-2000; 2000WO-US15446.
 XX 04-JUN-1999; 99US-0137693.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Yanagi M, Bukh J, Emerson SU, Purcell RH;
 XX WPI; 2001-061728/07.
 XX Nucleic acid molecule encoding human hepatitis C virus of genotype 2a
 PT for developing vaccines, for diagnosis of hepatitis C virus and in
 PT screening assays for identification of antiviral agents -
 XX Disclosure; Page 100-103; 167pp; English.
 CC AAC86645-48 represent chimeric cDNA clones of infectious Hepatitis C
 CC virus (HCV). In each clone the C, E1 and E2 genes are derived from HCV
 CC strain HC-J6CH genotype 2a. The p7 protein was derived from HCV strain
 CC HC-J6CH or HCV strain pCV-H77C, and the NS genes were derived from HCV
 CC strain pCV-H77C. Such HCV sequences are capable of expressing the virus
 CC when transfected into cells. The HCV protein is useful for assaying
 CC candidate antiviral agents for activity against HCV. Antibodies specific
 CC for HCV polypeptides are useful in prevention and treatment of diseases
 CC caused by HCV in animals, in particular humans. The HCV polypeptides
 CC serve as immunogens in the development of vaccines for preventing HCV
 CC in mammals or as antigens in diagnostic assays for detecting the
 CC presence of HCV in biological samples. The HCV polynucleotide is also
 CC useful for identifying cell lines capable of supporting the replication
 CC of HCV in vitro and to produce attenuated viral strains via passage in
 CC vitro or in vivo.
 XX Sequence 9611 BP; 1935 A; 2896 C; 2682 G; 2098 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 0 Length: 9611
 Score: 13049.00 Matches: 2395
 Percent Similarity: 87.40% Conservatives: 261
 Best Local Similarity: 78.81% Mismatches: 353
 Query Match: 80.95% Indels: 30
 DB: 22 Gaps: 4
 US-09-980-559-2 (1-3033) x AAC86645 (1-9611)
 QY 1 MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgProGln 20
 Db 342 ATGAGCACAAATCTCTAAACCTCAAAGAAAAACCAAGAAACACCAACCGTCCGCCACAA 401
 QY 21 AspValLysPheProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArg 40
 Db 402 GACGTTAAAGTTTCCGGCGCGCGCCAGATCGTTGGCGGAGTATATCTTGTTCGCCGCGAG 461
 QY 41 GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly 60
 Db 462 GGCCCGACGGTGGGTGTGGCGCGCGAGCAAGAAAGACTTCGAGCGGTGCCAGCACGTGGA 521
 QY 61 ArgArgGlnProIleProLysAspArgArgSerThrGlyLysSerTyrGlyLysProGly 80
 Db 522 AGCGCCAGCCCATCCCTAAAGATCGCGCTCCACTGGCAAAATCTTGGGGAACCAACGAGA 581
 QY 81 TyrProTyrProLeuTyrGlyAsnGluGlyLeuGlyTyrAlaGlyTyrLeuLeuSerPro 100
 Db 582 TACCCCTGGCCCTTATACGGGAATGAGGAGCTCGCGCTGGCGAGGATGGCTCTCTGTCCTCCC 641
 QY 101 ArgGlySerArgProSerTyrGlyProAsnAspProArgHisArgSerArgAsnValGly 120

[illegible]

4182	AAATATCAGAAACTGGAGTGGCGGACCGTTACACCGGGGAGCTCTATACATTACTTACTCCACTTAT	Ddb	4244
1301	GlyLysPheLeuAlaAspGlyGlyCysAlaGlyGlyValaTyArgPheIleIleCysAsp	Qy	1320
4242	GGCAAGTTTATCGCAGATGGAGCTGTGCAGCGGTGCCTATGACATCATATATCGGAC	Ddb	4301
1321	GluCysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGlnAla	Qy	1340
4302	GAATGCCATTCACTAGTGAGCGCTACTACCATCTTGGCATTTGGAACAGCTCTTGACCAAGCT	Ddb	4361
1341	GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProProGlySerValThr	Qy	1360
4362	GAGACCGCAGGCGCTCAGGCTAGTGGTGTGGCCACAGCCACGCCCTCCGGTACGGTGACA	Ddb	4421
1361	ThrProHisProAsnIleGluGluValAlaLeuGlyGlnGluGlyGluIleProPheTyR	Qy	1380
4422	ACTCCCAACAGTAACATAGAGAGGTGGCCCTTGGTCAAGGGGCGAGATCCCTTTTAT	Ddb	4481
1381	GlyArgAlaIleProLeuSerTyrlleLysGlyArgHisLeuIlePheCysHisSer	Qy	1400
4482	GGCAAAAGCTATTCCCTAGCTTTTCATCAAGGGGGGCGAGACACTTGATCTTTTGCCATTCA	Ddb	4541
1401	LysLysLysCysAspGluLeuAlaAlaLeuArgGlyMetGlyLeuAsnSerValAla	Qy	1420
4542	AAGAAGAAGTGGCAGAGCTCGCAGCGGGCCCTCCGGGGCAYGGGTGCTCAATCGCGTTGCA	Ddb	4601
1421	TyTyArgGlyLeuAspValSerValIleProThrGlnGlyAspValValValAla	Qy	1440
4602	TACTATAGGGGTCTCGACGCTCCCGTTATACCACTCAAGGAGAGCGTGGTGGTGTGCGCC	Ddb	4661
1441	ThrAspAlaLeuMetThrGlyTyThrGlyAspPheAspSerValIleAspCysAsnVal	Qy	1460
4662	ACTGATGCCCTAATGACTGGGTACACCGCGACTTTTGACTCTGTCATCGACTGTATATGT	Ddb	4721
1461	AlaValThrGlnValValAspPheSerLeuAspProThrPheThrIleThrGlnIle	Qy	1480
4722	GCAGTCTCTCAGATTGTGTGACTTCAGCCTTAGACCCCAACCTTCACCATCAACCTCAAAAC	Ddb	4781
1481	ValProGlnAspAlaValSerArgSerGlnArgGlyArgGlyArgGlyArgGlyArgLeu	Qy	1500
4782	GTCCCTCAGAGCGCTGTCTCCCGTAGTCAACGTAGAGGGGAGAACTGGAGGGGGCGATTG	Ddb	4841
1501	GlyIleTyArgTyArgValSerThrGlyGluArgAlaSerGlyMetPheAspSerValVal	Qy	1520
4842	GGCRRTTACAGGTATGTTTCGTCCAGYGARRGGCCGTCTGGGATGTTTCGACAGCGTAGTG	Ddb	4901
1521	LeuCysGluCysTyArgAspAlaGlyAlaAlaTyTyThrGluLeuThrProSerGluThrThr	Qy	1540
4902	CYCTCGAGTGTATGATCCCGGGCAGCGCTGTGTACGAGCTTTACACTCTCGAGACTAG	Ddb	4961
1541	ValArgLeuArgAlaTyRPhelanThrProGlyLeuProValCysGlnAspHisLeuGlu	Qy	1560
4962	GTGAGACTCCGGGCVATTTCACACGCGCCGGTTGCCCGTATGTCAAGACCAACCTCGAG	Ddb	5021
1561	PheTrpGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThr	Qy	1580
5022	TTCTGGGAACGGTCTTTACAGGTCTCAKCAKATTTACGCCCACTTCCTCTCCAGACG	Ddb	5081
1581	LysGlnSerGlyGluAsnPheAlaTyRLeuThrAlaTyRglnAlaThrValCysAlaArg	Qy	1600
5082	AAGCAAGGAGGAGAAACTTTGCTATCTAACGGCCCTACAGGCCACAGTATGCGCCAGG	Ddb	5141
1601	AlaLysAlaProProSerThrAspValMetTrpLysCysLeuThrArgLeuLysPro	Qy	1620
5142	GCAAAAGGCCCTCTCTTCGTGGGACGTGTGGAAAGTGTCCTAACTAGGCTCAAACT	Ddb	5201
1621	ThrLeuValGlyProThrProLeuTyArgLeuGlySerValThrAsnGluValThr	Qy	1640
5202	ACACTGACTGGTCCCACCCCTCTCTGTACCGCTTGGGTGCGGTGACCAATGAGGTYACC	Ddb	5261
1641	LeuThrHisProValThrLysTyrlleAlaThrCysMetGlnAlaAspLeuGluValMet	Qy	1660
5262	TTGAGCACCCTGACGAAATACATCGCAGCGTGCATGCAAGCTGACCTTGAGATCATG	Ddb	5321

1661	ThrSerThrTrpValLeuAlaGlyGlyValLeuAlaAlaValAlaAlaTyrCysLeuAla	1680
5322	ACAAGCTCATGGGTCTCTCGCGGGGGGTGTAGCCGCGTGGCAGCTTACTGCCTGGCG	5381
1681	ThrGlyCysValCysIleIleGlyArgLeuHisIleAsnGlnArgAlaValValAlaPro	1700
5382	ACTGCTGCATTTCCATCATTTGGCGCGCTACACCTGAAATGATCGGGTGGTTGTGRCCCY	5441
1701	AspLysGluValLeuTyrGluAlaPheAspGluMetGluGluCysAlaSerArgAlaAla	1720
5442	GACAGGAGATCTTATATAGGCTTTGTATGAGATGGAAGATGGCCCTCCAAAGCCGCC	5501
1721	LeuIleGluGluGlyGlnArgIleAlaGluMetLeuLysSerLysIleGlnIlyLeuLeu	1740
5502	CTCATTTAGGAAGGCGACGGATGCGGAGATGCTCAAAATCTAAGATACAAGCCTCTCTA	5561
1741	GlnGlnAlaSerLysGlnAlaGlnAspIleGlnProThrValGlnAlaSerTrpProLys	1760
5562	CAACAGGCGCAAGAGCAAGCTCAAGRCATRCAGCCAGCATATACAGTCAATCATGGCCAA	5621
1761	ValGluGlnPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeuAla	1780
5622	CTTGAAACAATTTTGGGCCCAACACATGTGGAACCTTCATCAGTGTATATACGTAAC	5681
1781	GlyLeuSerThrLeuProGlyAsnProAlaValAlaSerMetMetAlaPheSerAlaAla	1800
5682	GGACTCTCCACCTACCGGGAAATCTCTCAGTRGCATCAATGATGGCTTTTAGCGCCGCG	5741
1801	LeuThrSerProLeuSerThrSerThrThrIleLeuLeuAsnIleLeuGlyGlyTrpLeu	1820
5742	CTGACTAGCCCACTACCCACGAGCACCACTCCTCTTGAACATCATCGGAGGATGCTTG	5801
1821	AlaSerGlnIleAlaProProAlaGlyAlaThrGlyPheValValSerGlyLeuValGly	1840
5802	GCCTCYCAGATGGCCCCCTGCGCGAGCCACACYGCGCTTCGTGTGTCAGTGGTCTAG	5861
1841	AlaAlaValGlySerIleGlyLeuGlyLysValLeuValAspIleLeuAlaGlyTyrGly	1860
5862	CGGCGCTCGAGACATAGGCTGGGTAAAGATACGTGTGGACGTTTTGGCCGGGTACGGC	5921
1861	AlaGlyIleSerGlyAlaLeuValAlaPheLysIleMetSerGlyGluLysProSerMet	1880
5922	GCAGGCATTTCAGGGGCCCTCGTAGCTTTTAAGATCATGAGCGCGGAGAACCCACGGTA	5981
1881	GluAspValValAsnLeuLeuProGlyIleLeuSerProGlyAlaLeuValValGlyVal	1900
5982	GAAGACGTTGTGAATCTCCTCGCTGCTATYCTGTCTCCTGTGTGGTGTGGTAGTGGAG	6041
1901	IleCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrpMet	1920
6042	ATCTGTGAGCAATYCTGCGCGCCACCTCGTCTAGGAGAGGGKCCGGTCCAGTGGATG	6101
1921	AsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaProThrHisTyrValThr	1940
6102	AACGACTGATCGCTTCCGCTCCAGGGGAAACACACGTTGCCCCCTACCCACTACGTGGTG	6161
1941	GluSerAspAlaSerGlnArgValThrGlnLeuLeuGlySerLeuThrIleThrSerLeu	1960
6162	GAGTCTGACGCTTACACAGCGTGTACGCAAGGTGCTGAGTTTCACTTACAAATTTACCA	6221
1961	LeuArgArgLeuHisAsnTrpIleThrGluAspCysProIleProCysGlyGlySerTrp	1980
6222	CTTAGGAGACTACATGCTCGATCACTGAAAGATGCCCAATGCCCATGCTCGGGGTCCTGG	6281
1981	LeuArgAspValTrpAspTrpValCysThrIleLeuThrAspPheLysAsnTrpLeuThr	2000
6282	CTCCAGGACATTTGGGATTTGGTTTGTTCATCTCTTCACAGACTTVAATAAACTGGTGTCT	6341
2001	SerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTyrLys	2020
6342	TCAAAATATCTCCCAAGATCCCGGATCTCCCTTTATCTCTTGCAGAGGAGATACAAAG	6401

561 GlyTyrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer 580
 2022 GGGTTCACTAAGACATCGGTGACACCTTGGCGCATTTAGGAAGACTACACAGCACT 2081
 581 ThrAspLeuLeuCysProThrAspCysPheArgGlyHisProAspThrThrTyrLeuLys 600
 2082 CTCGATTTATTTGGCCCCACAGACTGTTTGGGAAGCACCAGATCTACTTCTTAAG 2141
 601 CysGlySerGlyProThrLeuThrProArgCysLeuLeuAspTyrProTyrArgLeuTyr 620
 2142 TGTGGAGAGGGGCTTGTGTTAACTCCAGGTGCTGTGTAGCTACCTTATAGTYGTGG 2201
 621 HisTyrProCysThrValAsnTyrThrIlePheLysIleArgMetTyrValGlyGlyVal 640
 2202 CATTATCCGTGCACCTGTAACTTCAACCTTAAAGCGCGGATGTATGTAGGAGGGTG 2261
 641 GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp 660
 2262 GAGCATCGATTTCTCCGAGCATGCACTTCACTGCGGGAGATCGCTGAGACTTGAAGAT 2321
 661 ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGlnThrAlaIleLeuPro 680
 2322 AGGATAGGGGCGAGCAGAGTCCACTGCTGCATTTCCACTACTGAGTGGGGGTGTCCTCA 2381
 681 CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisLeuHisGlnAsnIle 700
 2382 TGCTCTCTCTGACCTACAGCAGCATCTCACTGGGCTATTGACCTCCACCAAAACATC 2441
 701 ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrLysTyrIleValArgTyr 720
 2442 GTGGACGTGAGTACCTTTCGCGACTTCTTCGCGCTCTGCAAGATACATCGTGAAGTGG 2501
 721 GluTyrValIleLeuLeuPheLeuLeuAlaAspAlaArgValCysAlaCysLeuTyr 740
 2502 GAGTGGGTGATCTCTCTTCTTGTGTGGAGACCCAGGACTGTGTGATGCCITTTGG 2561
 741 MetLeuIleLeuLeuGlyGlnAlaGluAlaLeuGluLysLeuValIleLeuHisAla 760
 2562 ATGCTCAWCACTACTGGGCGAAGCGGAGCGGCTTGAAGCTCATCTTGCATCTCC 2621
 761 AlaSerAlaAlaSerCysAsnGlyPheLeuTyrPheValIlePhePheValAlaAlaTyr 780
 2622 GCTAGYCTGCTAGTGCATATGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2681
 781 TyrIleLysGlyArgValValProLeuAlaThrThrThrSerLeuThrGlyLeuTyrSerPhe 800
 2682 TACTTAAAGGCGAGGGTGGTCCCGTGGCCAGCTACTCTGTCTGCTGCTGCTGCTGCT 2741
 801 SerLeuLeuLeuAlaLeuProGlnGlnAlaTyrAlaTyrAspAlaSerValHisGly 820
 2742 CTCCTCTCTAGTCTGGCTTTACACAGCAGGCTTATGCTTGGAGCTGTGGAACAGGG 2801
 821 GlnIleGlyAlaAlaLeuLeuValMetIleThrLeuPheThrLeuThrProGlyTyrLys 840
 2802 GAACCTGGGCTGGCCATATTAGTAATATATATATATATATATATATATATATATAT 2861
 841 ThrLeuLeuSerArgPheLeuTyrProLeuTyrProLeuTyrProLeuTyrProLeuTyr 860
 2862 ATCTCTGAGCCGCTTCAAGTGTGGTGTGCTGCTTACATGCTGCTGCTGCTGCTGCTG 2921
 861 ValGlnGluTyrAlaProProMetGlnValArgGlyGlyArgAspGlyIleIleTyrAla 880
 2922 ATTCAGCAATGGGTTCCTCCCTCGAGGTCCGAGGGGGGCTGAGCGGATCATCTGGGTG 2981
 881 ValAlaIlePheTyrProGlyValValPheAspIleThrLysThrLeuLeuAlaValLeu 900
 2982 GCTGTCATCTACACCCAGCCTTGTGTGTGGAGTTCAGAAATGGTGTGTGTGTGTGTGT 3041
 901 GlyProAlaTyrLeuLeuLysGlyAlaLeuThrArgValProTyrPheValArgAlaHis 920
 3042 GGGGCTGCTTACCTCTTARAGCGCTCTGCTACGGATACCTGCTTGTGTGTGTGTGTGT 3101
 921 AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyGlyArgTyrValGlnMet 940

3102 GCTTGTGTACGAGTGTACCTGTGTGAACACCTCGRGGGCTAGGTACATCCAGATG 3161
 941 AlaLeuLeuAlaLeuGlyArgThrThrGlyThrTyrIleTyrAspHisLeuThrProMet 960
 3162 CTGTGTATCATCATAGGAGATGGACCGGCTTACATCATACACACCTCTCCCTTTA 3221
 961 SerAspTyrAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIlePhe 980
 3222 TCAACTTGGGCGGCCAGGGTTTTCGGGACCTTGGCAATCGCGTGGAGCTGTGGTGTTC 3281
 981 SerProMetGluLysIleValIleValTyrGlyAlaGluThrAlaAlaCysGlyAspIle 1000
 3282 AGCCCAATGAGAGAAGGTCATTTGTGGGGGCTGAGACAGTGGCGTGTGGAGACATC 3341
 1001 LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp 1020
 3342 CTGATGGGCTCCCGTCTCCCGAGGCTAGGTAGGARGTTCTGCTCGGCTGCGGAC 3401
 1021 GlyTyrThrSerLysGlyTyrSerLeuLeuAlaProIleThrAlaTyrAlaGlnThr 1040
 3402 GGCTACACCTCCAGGGGTGAACCTCTCTAGCTCCCATTTACTGCTTACACTCAGCAACT 3461
 1041 ArgGlyLeuLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGluGlnAla 1060
 3462 CGTGGTCTCTGGTGTCTATCGTGTGCTAGCTTAAACGGGCGGCAAAATGAGCAGGCT 3521
 1061 GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly 1080
 3522 GGGGAGGTCAGGCTCTGCTCCGCTCACAAACCTCTTGGGGACATCTCATTTCCGGC 3581
 1081 ValLeuTyrThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro 1100
 3582 GTCTCTGACAGTATATACGGGCTGTGTATAGACCTTGGCGGCGCCCAAGGAGCA 3641
 1101 ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTyrProSerProGly 1120
 3642 GTCACTCAGATGTACACAGCGAGAGGGGAGCTCTGTTGGGATGGCTAGTCTCCCGGG 3701
 1121 ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArgAsn 1140
 3702 ACTAAGTCATGAGCCCTTGTACTCGGGGCGCTAGACCTCTACTGTGTGCTACCCGAAAC 3761
 1141 AlaAspValIleProAlaArgArgGlyAspLysArgGlyAlaLeuLeuSerProArg 1160
 3762 GCTGATGTCTCCGTCGGAGAAAGATGACCGAGGGGTGCTACTTCTCGCAAGG 3821
 1161 ProLeuSerThrLeuLysGlySerSerGlyGlyProValLeuCysProArgGlyHisAla 1180
 3822 CCCCTCTCAACCTCAAGGATCATCGGAGGGCGCTGCTCTGCTCAGGGGACACGCC 3881
 1181 ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle 1200
 3882 GTGGGCTGTGTACAGCGCGGCTGTGTCAGGGGTGTAGCCAAATCTATTGATCTTATC 3941
 1201 ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro 1220
 3942 CCGCTCGAATCATCTGATTCGCCACAGCGGCGCCAGTCTTCTCTGACACAGTTCGCGG 4001
 1221 ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLys 1240
 4002 CCAGCTGTGCGCCAGTCTTACCAGGTGGGTACTTGCACGCGCACCAAGCAGCGGGAAG 4061
 1241 SerThrLysValProValAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnPro 1260
 4062 AGCACAAGTCTCCCGGTATGCGAGTACGAGGGGTATTAAGTACTCTGCTACTTAATCCC 4121
 1261 SerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerLysAlaHisGlyIleAsnPro 1280
 4122 TCTGTGCGGCGCACACTTGTGTTGGGGCTACATGTCCAAAGCCACGCGGATCAACCCCT 4181
 1281 AsnIleArgThrGlyValArgThrValThrThrGlyAlaProIleThrTyrSerThrTyr 1300

XX Nakamura T, Okamoto H;
 XX WPI; 1993-087166/11.
 XX N-PSDB; AAR33539, AAR33214.
 DR Polynucleotide(s), polypeptide(s) and antibodies of NANBH virus -
 XX useful for detecting NANBH, as a vaccine and for screening blood
 XX samples
 XX Claim 7; Page 59-64; 93pp; English.
 XX RNA was isolated from the plasma of human patients positive for
 CC NANBH virus (strain HC-J8) and was subjected to reverse transcription
 CC to produce cDNA. The resulting cDNA was amplified by PCR, and
 CC nucleic acid sequences determined by analysis of clones obtd. by PCR
 CC amplification (42 clones in total). The NANBH HC-J8 degenerate genome
 CC was found to contain a degenerate open reading frame encoding
 CC polypeptide precursors of 3033 amino acid residues.
 CC See also AAO38172-220.
 XX (Updated on 25-MAR-2003 to correct PN field.)
 XX Sequence 9511 BP; 2020 A; 2709 C; 2539 G; 2108 T; 135 other;
 SQ

Alignment Scores:
 Pred. No.: 0 Length: 9511
 Score: 13810.00 Matches: 2523
 Percent Similarity: 90.93% Conservative: 235
 Best Local Similarity: 83.18% Mismatches: 275
 Query Match: 85.67% Indels: 0
 DB: 14 Gaps: 0

US-09-980-559-2 (1-3033) x AAO38221 (1-9511)

QY	1	MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGln	20
DB	342	ATGAGCACAATTCCTAAACCTCAAGAAGAAACCAAAAGAAACACAAACCGCGCCACAG	401
QY	21	AspValLysPheProGlyGlyGlyGlnIleValGlyGlyValTyrlLeuLeuProArgArg	40
DB	402	GACGTTAAGTTCCCGGGTGGCGGTTCAGATGTTGGCGGAGTTTACTTTGTCGCCGCAGG	461
QY	41	GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly	60
DB	462	GGCCCCAGTTGGGTGTGGCGGCACAGGAAGAACTTCYGAGCGATCCCGACCGCGTGA	521
QY	61	ArgArgGlnProIleProLysAspArgArgSerThrGlyLysSerTrpGlyLysProGly	80
DB	522	CGAGCCAGCCCATCCCCGAAAGATCGCGCTCCACCGCGAAGTCTCTGGGGAAGACGAG	581
QY	81	TyrProTrpProLeuTyrGlyAsnGluGlyLeuGlyTyrAlaGlyTrpLeuLeuSerPro	100
DB	582	TATCCTTGGCCCCCTGTACGGAAACAGAGGTTTGGCTGGCGCGGGTTGGCTTCCTGTC	641
QY	101	ArgGlySerArgProSerTrpGlyProAsnAspProArgHisArgSerArgAsnValGly	120
DB	642	CGCGGGTCTGTCCTACTTTGGGGCCCCACCGACCCCCGGCATAGATCAGCAATTTGGG	701
QY	121	LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProValVal	140
DB	702	AGAGTCATCGATACCATTCAGTGTGTGGTTTTTGGCCGACCTCATGGGGTACATCCCT	761
QY	141	GlyAlaProLeuGlyGlyValAlaArgAlaLeuAlaHisGlyValArgValLeuGluAsp	160
DB	762	GGCGCCCCGGTYGGAGGCGTCCGACAGCTCTGGCACACGGTGTTAGGGTCTCTGGAG	821
QY	161	GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla	180
DB	822	GGGATAAATACGACACAGGAATTAACCGGTGTCTCTTTTCTATCTTTTGTCTGTCT	881
QY	181	LeuLeuSerCysIleThrThrProValSerAlaAlaGluValLysAsnIleSerThrGly	200
DB	882	CTTCTGTATCCGTCACARTGCCAGTCTCTGCACTGGAGTCAAGAACATYAGTTCTGAC	941

QY	2461	GlulysLeuProIleAsnProLeuSerAsnSerLeuLeuArgTyrHisLenLysValTyr	2480	QY	2821	ArgArgTyrTyrLeuThrArgAspProThrProIleAlaArgAlaThrGlu	2840
DB	7721	GAAAGTGGCAATCAACCTTTGAGTAACTCGCTGTGGATACCAATAACAGGTGTAC	7780	DB	8801	CGCCGAGATACCTACCTGACAGAGAGACCAACCACTCCACCGCGCTGCTGGAA	8860
QY	2481	CysThrThrThrLysSerAlaSerLeuArgAlaLysValThrPheAspArgMetGln	2500	QY	2841	ThrValArgHisSerProValAsnSerTrpLeuGlyAsnIleIleGlnTyrAlaProThr	2860
DB	7781	TGTACAACTCAAGAGCGCTCACAGAGCGCTTAAAGAGTAACTTTTGACAGGAGCAA	7840	DB	8861	ACAGTTAGACATCTCCCTATCAATCATGGCTGGGAAACATCATCCAGTATGCTCAACC	8920
QY	2501	ValLeuAspSerTyrTyrAspSerValLeuLysAspIleLysLeuAlaSerLysVal	2520	QY	2861	IleTTrpAlaArgMetValLeuMetThrHisPheSerIleLeuMetAlaGlnAspThr	2880
DB	7841	GTGCTCGAGCCCATTAATGACTCAGTCTTAAAGGACATCAAGTAGCGCTTCAAGGTC	7900	DB	8921	ATATGGTTCGATGGTCTTAATGACACACTTCTTCTCCATCTCATGGTCCAGACACC	8980
QY	2521	ThrAlaArgLeuLeuThrMetGluAlaCysGlnLeuThrProProHisSerAlaArg	2540	QY	2881	LeuAspGlnAsnLeuAsnPheGluMetTyrGlyAlaValTyrSerValSerProLeuAsp	2900
DB	7901	AGCGCAAGGCTCTCACCTTGGAGAGCGCTGCCAGTTGACTCCACCCCATCTGCAAGA	7960	DB	8981	CTGGACCAAGACTCAACTTTGAGATGTATGGATCATCTCCGTGAATCCTTTGGAC	9040
QY	2541	SerLysTyrGlyPheGlyAlaLysGluValArgSerLeuSerGlyArgAlaValAsnHis	2560	QY	2901	LeuProAlaIleIleGluArgLeuHisGlyLeuAspAlaPheSerLeuHisThrTyrThr	2920
DB	7961	TCCAAGTATGGATTTCGGGGCAAGGAGTCCGAGCTTGTCCGGAGGGCGGTTAACCCAC	8020	DB	9041	CTTCCAGCCATTAATTGAGAGGTTACACGGCTTGACGCTTCTTCTATGCACATACTCT	9100
QY	2561	IleLysSerValTrpLysAspLeuLeuGluAspSerGluThrProIleProThrIle	2580	QY	2921	ProHisGluLeuThrArgValAlaSerAlaLeuArgLysLeuGlyAlaProProLeuArg	2940
DB	8021	ATCAAGTCCGTGGAGAGACCTCTCTGAAGACCCCAACACCAATTTCCCAACCAATC	8080	DB	9101	CACCACGAACCTGACGCGGTGGCTTCAGCCCTCAGAAAACTTGGGGCGCCACCCCTCAGG	9160
QY	2581	MetAlaLysAsnGluValPheCysValAspProThrLysGlyLysLysAlaAlaArg	2600	QY	2941	AlaTrpLysSerArgAlaArgAlaValArgAlaSerLeuIleSerArgGlyGlyArgAla	2960
DB	8081	ATGGCCAAAATAATGAGGTGTTCTGGTGGACCCCGCCCAAGGGGGTGAAGAACCCAGCTCG	8140	DB	9161	GTGTGAAGAGTGGGCTCGCGCAGTCAGGGCTCCCTCATCTCCGCTGGAGGAAAGCG	9220
QY	2601	LeuIleValTyrProAspLeuGlyValArgValCysGluLysMetAlaLeuTyrAspIle	2620	QY	2961	AlaValCysGlyArgTyrLeuPheAsnTrpAlaValLysThrLysLeuLysLeuThrPro	2980
DB	8141	CTCATCGTTTACCTGACCTCGCGCTCGCGTCTCGGAGAAATGGCCCTCTATGACATT	8200	DB	9221	GCGGTTTGGCGCGATATCTTCAATGGCGGTGAGACCCGCTCATCTCCGCTGGAGGAAAGCG	9280
QY	2621	ThrGlnLysLeuProGlnAlaValMetGlyAlaSerTyrGlyPheGlnTyrSerProAla	2640	QY	2981	LeuProGluAlaArgLeuLeuAspLeuSerSerTrpPheThrValGlyAlaGlyGlyGly	3000
DB	8201	ACCAAAAGCTTCTCAGGCGGTAATGGAGCTTCTTATGGCTTCCAGTACTCCCTGCC	8260	DB	9281	TTGCGGAGGCGCGCTTACTGGACTTATCCAGTTGTTTCCCGCTCGCGCGCGCGGGCG	9340
QY	2641	GlnArgValGluPheLeuLeuLysAlaTrpAlaGluLysLysAspProMetGlyPheSer	2660	QY	3001	AspIleTyrHisSerValSerArgAlaArgProArgLeuLeuPheGlyLeuLeuLeu	3020
DB	8261	CAACGGGTGAGTATCTCTTGAAGCATGGCGGAAAGAGAACCCCTGCGTTTTCG	8320	DB	9341	GACATTTTTCACAGCGTGTGCGCGCGCGCGCGCGCGCTCATCTTCTCGGCTACTCCTA	9400
QY	2661	TyrAspThrArgCysPheAspSerThrValThrGluArgAspIleArgThrGluGluSer	2680	QY	3021	LeuPheValGlyValGlyLeuPheLeuLeuProAlaArg	3033
DB	8321	TATGATACCCGATGCTTTCGACTCAACCGTCACTGAGAGAGACATCAGGACCGAGGTCC	8380	DB	9401	CTTTTCGTAGGGTGAAGCTTCTTCTTACTCCCGCTCGG	9439
QY	2681	IleTyrArgAlaCysSerLeuProGluAlaHisThrAlaIleHisSerLeuThrGlu	2700	DB	9401	CTTTTCGTAGGGTGAAGCTTCTTCTTACTCCCGCTCGG	9439
DB	8381	ATATACCAGGCTGCTCTCCCTGCGGAGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT	8440	DB	9401	CTTTTCGTAGGGTGAAGCTTCTTCTTACTCCCGCTCGG	9439
QY	2701	ArgLeuTyrValGlyGlyProMetPheAsnSerLysGlyGlnThrCysGlyTyrArgArg	2720	DB	9401	CTTTTCGTAGGGTGAAGCTTCTTCTTACTCCCGCTCGG	9439
DB	8441	AGACTTTACGTAGAGGGCCCATGTTCAACAGCAAGGGTCAACCTCGGCTTACAGAGCT	8500	DB	9401	CTTTTCGTAGGGTGAAGCTTCTTCTTACTCCCGCTCGG	9439
QY	2721	CysArgAlaSerGlyValLeuThrThrSerMetGlyAsnThrIleThrCysTyrValLys	2740	DB	9401	CTTTTCGTAGGGTGAAGCTTCTTCTTACTCCCGCTCGG	9439
DB	8501	TGCGCGCGCGCGGCTGCTAAACCTAGCATGGGTAAACCATCATGCTATGTGAAA	8560	DB	9401	CTTTTCGTAGGGTGAAGCTTCTTCTTACTCCCGCTCGG	9439
QY	2741	AlaLeuAlaAlaCysLysAlaAlaGlyIleIleAlaProThrMetLeuValCysGlyAsp	2760	DB	9401	CTTTTCGTAGGGTGAAGCTTCTTCTTACTCCCGCTCGG	9439
DB	8561	GCCCTAGCGGCTGCAAGGCTGCGGGGATAGTTGCGGCCCAATGCTGGTATGCGGGAT	8620	DB	9401	CTTTTCGTAGGGTGAAGCTTCTTCTTACTCCCGCTCGG	9439
QY	2761	AspLeuValValIleSerGluSerGlnGlyThrGluGluAspGluArgAsnLeuArgAla	2780	DB	9401	CTTTTCGTAGGGTGAAGCTTCTTCTTACTCCCGCTCGG	9439
DB	8621	GACCTAGTATGCTCATCTCAGAAAGCCAGGGGCTGAGGAGGAGGAGGAGGAGGAGGAG	8680	DB	9401	CTTTTCGTAGGGTGAAGCTTCTTCTTACTCCCGCTCGG	9439
QY	2781	PheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProArgProGluTyr	2800	DB	9401	CTTTTCGTAGGGTGAAGCTTCTTCTTACTCCCGCTCGG	9439
DB	8681	TTACGGAGGCGCATGACAGGTACTTGCCTCTCTGCTGATCCCCCAGACCGCAATAT	8740	DB	9401	CTTTTCGTAGGGTGAAGCTTCTTCTTACTCCCGCTCGG	9439
QY	2801	AspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaLeuGlyProGlnGly	2820	DB	9401	CTTTTCGTAGGGTGAAGCTTCTTCTTACTCCCGCTCGG	9439
DB	8741	GACCTGGAGCTTAATAACATCTGTTCTCTCAATGTGTGTGGGCTTGGGGCGCGGGC	8800	DB	9401	CTTTTCGTAGGGTGAAGCTTCTTCTTACTCCCGCTCGG	9439

5501	DB	CTCATCAAGAGGGGAGGGGATAGCCGAGATGTTGAAGTCCAAAGGCTTGGCTG	5560	DB	AACTAACAAGCCGCATCTGGAGGGTGGCGCTCGGAGTACGCGAGGTGACGACGACAT
1741	QY	GlnGlnAlaSerLysGlnAlaGlnAspIleGlnProThrValGlnAlaSerProLys	2101	QY	GlySerTyrHisTyrIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeu
5561	DB	CAGCAGGCTCTTAGCAGGCCCAAGGACATACACCCGCTATGCGGCTTCATGGCCAAA	6641	DB	GGGTGCTACTCTATGTAAACAGACTGACCATGACATCTGAAATCTCTGGCAACTA
1761	QY	ValGlnGlnPheTrpAlaLysHisMetTrpAsnPhelileSerGlyIleGlnTyrLeuAla	2121	QY	ProSerProGluPhePheSerTrpValAspGlyValGlnIleHisArgPheAlaProThr
5621	DB	GTGGAACAATTTGGGCCAGACACATGTGAACCTTCATTAGCGGCATCCAACTCCGA	6701	DB	CCTTCTCCAGAGTCTTCTCGGTGGAGCGGTGCAGATCCATAGGTTGCAACCA
1781	QY	GlyLeuSerThrLeuProGlyAsnProAlaValAlaSerMetMetAlaPheSerAlaAla	2141	QY	ProLysProPhePheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal
5681	DB	GCATTCTCAACACTGCACAGGAACCCCGCGTTCCTTCATGATGCAATTCAGTCCGCC	6761	DB	CCAAAGCCGTTTTTCGGGATGAGGTCTCGTTCTCGGTTGGGCTTAAATCTTATGCTGC
1801	QY	LeuThrSerProLeuSerThrThrThrIleLeuLeuAsnIleLeuGlyTyrProLeu	2161	QY	GlySerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThr
5741	DB	CTACACAGTCCGTTGTCAGCACATACCATCTTCTCAACATCATGGGAGGCTGGTTA	6821	DB	GGGTCCCAAGCTTCCCTGTGAACCTGAGCCCGGAGCGAGCGTATGAGGTCCATGCTAACA
1821	QY	AlaSerGlnIleAlaProAlaGlyAlaThrGlyPheValValSerGlyLeuValGly	2181	QY	AspProSerHisIleThrAlaGluThrAlaAlaArgArgLeuAlaArgGlySerProPro
5801	DB	CGCTCCAGATCGCACACCCCGGGGGCCACCGGCTTTGTCAGTGGCCCTGGTGGGG	6881	DB	GATCCGCCACATCACGGCGGAGACTGGCGCGCGCGCTTGGCAGCGGATCACCTCCA
1841	QY	AlaAlaValGlySerIleGlyLeuGlyLysValLeuValAspIleLeuAlaGlyTyrGly	2201	QY	SerGluAlaSerSerAlaSerGlnLeuSerAlaProSerSerLeuAlaThrCysThr
5861	DB	GCTGCCGTGGCAGCATAGCCCTGGGTAAAGTCTGCTGGACATCTCTGGCAGGATATG	6941	DB	TCTGAGCGGAGCTCTCTGAGCCAGCTATCAGCACCGCTCGCTGCGGCGCACCTGCACC
1861	QY	AlaGlyIleSerGlyAlaLeuAlaPheLysIleMetSerGlyLysProSerMet	2221	QY	ThrHisGlyLysAlaTyrAspValAspMetValAspAlaAsnLeuPheMetGlyGlyAsp
5921	DB	GGGGGATTTCCGGGGCCCTCGTCGATTCAGATCATCTGTGGCGAGAGCCCTCTATG	7001	DB	ACCCACAGCACACCTATGACGTGACATGGTGCATGCCACCTGCTCATGAGGCGCGGT
1881	QY	GluAspValValAsnLeuLeuProGlyIleLeuSerProGlyAlaLeuValGlyVal	2241	QY	ValThrArgIleGlySerGlySerLysValValValLeuAspSerLeuAspProMetVal
5981	DB	GAAGATGTCACTCACTACTGCTGGGATCCTGTCTCCGGGAGCCCTGGTGGGGGTC	7061	DB	GTGGCTCAGACAGAGCCTGAGTCCAGGGTCCCGTCTTGACTTCTCGAGCAATGGCC
1901	QY	IleCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrpMet	2261	QY	GluGluArgSerAspLeuGluProSerIleProSerGlyTyrMetLeuProLysLysArg
6041	DB	ATCTGCCGCGCCATCTCGCGCGCCACGCTGGGACCGGGGAGGGCGGCTCCAAATGGATG	7121	DB	GAGGAAGAGAGAGCAGCTTGGAGCCCTCAATACCATCGGAGTGTCTCCCGAGCGCGG
1921	QY	AsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaProThrHisTyrValThr	2281	QY	PheProAlaLeuProAlaTrpAlaArgProAspTyrAsnProProLeuValGluSer
6101	DB	AACAGGCTTATGCTTTGCTTCCAGAGGAACACGCTCGCCCTACTCACTACGTGACG	7181	DB	TTTCCAGGGGCTTACCGGCTTGGGACGCGCTGACTACACCCGCGCTCTGTTGGAATCG
1941	QY	GluSerAspAlaSerGlnArgValThrGlnLeuGlySerLeuThrIleThrSerLeu	2301	QY	TrpLysArgProAspTyrGlnProAlaThrValAlaGlyCysAlaLeuProProArg
6161	DB	GAGTGGATGCTGCGAGCGGTGACCAACTACTTGGCTCTCTACTATATACACGCTA	7241	DB	TCGAGGAGGCCAGATTACCAACCGCCACCGTGTGTGTGTGTGTCTCTCCCGCCCAAG
1961	QY	LeuArgArgLeuHisAsnTrpIleThrGluAspCysProIleProCysGlyGlySerTrp	2321	QY	LysThrProThrProProArgArgArgThrValGlyLeuSerGluAspSerIle
6221	DB	CTCAGAAGACTCCACAATTTGGATACTGAGGACTGCGCCCATCCCATGCTCCGGATCCTCG	7301	DB	AAGGCCCGGACGCTCCCGCAAGGAGAGCGCGGACAGTGGGTCTGAGCGAGACACATA
1981	QY	LeuArgAspValTrpAspTrpValCysThrIleLeuThrAspPheLysAsnTrpLeuThr	2341	QY	GlyAspAlaLeuGlnGlnLeuAlaIleLysSerPheGlyGlnProProProSerGlyAsp
6281	DB	CTCCGCGAGCTGTGGGACTGGGTTCACCATCTTGACAGACTTCAAAAATTTGGCTGACC	7361	DB	TCAGAAGCCCTCCAGCAACTGGCCATCAAGACTTTTGGCAGCCCTCCCGGCGGTGAT
2001	QY	SerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTyrLys	2361	QY	SerGlyLeuSerThrGlyValAlaGlyAlaAspSerGlySerGlnThrProProAspGlu
6341	DB	TCFAAATTTGTTCCCAAGCTGCCCGCCCTCCCTTCTCTTGTCAAAAGGGGTACAAG	7421	DB	GCAGGCTGCTCCACGGGGGCGCGCGCGCGGATCCGGGGTCCGAGCTCCCTGCTGAG
2021	QY	GlyValTrpAlaGlyThrGlyIleMetThrThrArgCysProCysGlyAlaAsnIleSer	2381	QY	LeuAlaLeuSerGluThrGlySerIleSerSerMetProProLeuGluGlyGluLeuGly
6401	DB	GGTGTGTGGCGCGGCACTGGCATCATGACACGCGCTGCCCTTGGCGCGCCCAACTCTCT	7481	DB	CCGGCCCCCTCAGAGACAGGTTCCGCTCTCTATGCCCCCTCCGAGGGGAGCGCTGGA
2041	QY	GlyAsnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTrp	2401	QY	AspProAspLeuGluProGluGlnValGluProGlnProProProGlnGlyGlyValAla
6461	DB	GGCAATGTCCGCTGGGCTCTATGAGGATCACAGGCGCTTAAACCTGATGAACCTGG	7541	DB	GATCCGACCTGGAGTCTGATCAGGTAGAGCTTCAACCTTCCCGGCGGGGGGGGTA
2061	QY	GlnGlyThrPheProIleAsnCysTyrThrGluGlnCysValProLysProAlaPro	2421	QY	AlaProGlySerAspSerGlySerTrpSerThrCysSerGluGluAspAspSerValVal
6521	DB	CAGGGAGACCTTCTCTATCAATGCTACAGGGAGGCGAGTGGCGCGGAAACCCCGCAG	7601	DB	GCTCCGGTTCGGGCTCGGGCTTGGTCTTACTTCTCGAGGAGGAGGACGATACCGG
2081	QY	AsnPheLysValAlaIleTrpArgValAlaAlaSerGlyTyrAlaGluValThrGlnHis	2441	QY	CysCysSerMetSerTyrSerTrpThrGlyAlaLeuIleThrProCysSerProGluGlu
			7661	DB	TGCTGCTCATGTCATCTCTCTGGACCGGGGCTCTAAATACTCTCTGTAGCCCGGAGAG

Qy	1001	LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp	1020	Qy	1361	ThrProHisProAsnIleGluGluValAlaLeuGlyGlnGluGlyGluLeuProPheTyr	1380
Db	3341	CTACATGGACTCCCGTGTCCCGGACTCGGCAGAGATCTCTTCGGCCCGACTGAT	3400	Db	4421	ACCCCCCATCCGATATAGAGAGGTAGCCCTCGGCGGAGGTGAGATCCCTCTAT	4480
Qy	1021	GlyTyrThrSerLysGlyTyrSerLeuLeuAlaProIleThrAlaGlnGlnThr	1040	Qy	1381	GlyArgAlaIleProLeuSerTyrIleLysGlyArgHisLeuLeuPheCysHisSer	1400
Db	3401	GGCTACACCTCCCAAGGGGTGAAGCTCTTGTCTCCCATCACTCTTATGTCAGCAACA	3460	Db	4481	GGAGGGCGATTCCTTCATCCATCAAGGGAGGAGACCTGATTTCTGCCACTCA	4540
Qy	1041	ArgGlyLeuLeuGlyThrIleValSerMetThrGlyArgAspLysThrGluGlnAla	1060	Qy	1401	LysLysLysCysAspGluLeuAlaAlaLeuLeuArgGlyMetGlyLeuAsnSerValAla	1420
Db	3461	CGAGGCTCTCGGCGCATAGTGGTGAATGACCGGGCGTGCACAGGACAGAGGCC	3520	Db	4541	AGAAGAAAGTGCACAGCTCCGCGGCGCTTCGGGCGATGGGCTTGAATCCGCTGCA	4600
Qy	1061	GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly	1080	Qy	1421	TyrTyrArgGlyLeuAspValSerValIleProThrGlnGlyAspValValValAla	1440
Db	3521	GGGAAGTCCAAATCTCTCCAGTCTCTCACTCTCTCTCGAACAACCATCTCGGG	3580	Db	4601	TACTATAGAGGTGGACGTCTCCATAATACCACTCAGGAGATGTGTGTGCTGCTGCGC	4660
Qy	1081	ValLeuTyrThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro	1100	Qy	1441	ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnVal	1460
Db	3581	GTTTGTGGACTGTTTACACCGAGCTGGCAACAAGACTCTAGCCGGCTTACGGGTCCG	3640	Db	4661	ACCGAGCGCTCATGACGGGTACACTGGAGACTTTTGACTCCGTGATCGACTGCAATGA	4720
Qy	1101	ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTyrProSerProProGly	1120	Qy	1461	AlaValThrGlnValValAspPheSerLeuAspProThrPheThrIleThrThrGlnIle	1480
Db	3641	GTCAAGCAGATGACTCGAGTGTGAGGGGACTTGGTAGGCTGGCCAGCCCGCTGGG	3700	Db	4721	GGGTCAACCAAGCTGTGCACCTTCAGCTTGACCCCACTTCACTATACACACAGACT	4780
Qy	1121	ThrLysSerLeuGluProCysThrCysGlyValAlaValAspLeuTyrLeuValThrArgAsn	1140	Qy	1481	ValProGlnAspAlaValSerArgSerGlnArgGlyArgThrGlyArgGlyArgLeu	1500
Db	3701	ACCAAGTCTTTGGAGCGTGAAGTGTGGAGCGGTGCACCTATCTGTGTGTCAGCGGAAC	3760	Db	4781	GTCCCAACAAGCGTGTCTCACGAGTCAGCGCGCGCGCACAGGTAGAGAGACAG	4840
Qy	1141	AlaAspValIleProAlaArgArgGlyAspLysArgGlyAlaLeuLeuSerProArg	1160	Qy	1501	GlyIleTyrArgTyrValSerThrGlyGluArgAlaSerGlyMetPheAspSerValVal	1520
Db	3761	GCTGATGTCATCCCGGCTCGGAGACGCGGGGCAAGCGGGGAGCATTTGCTCTCCCGAGA	3820	Db	4841	GGCAGCTTATAGGTATGTTTCCACTGGTGAACGAGCCTCAGGAATGTTTGCAGAGTAGTG	4900
Qy	1161	ProLeuSerThrLeuLysGlySerSerGlyGlyProValLeuLysCysProArgGlyHisAla	1180	Qy	1521	LeuCysGluCysTyrAspAlaGlyAlaAlaTyrTyrGluLeuThrProSerGluThrThr	1540
Db	3821	CCCATTTGACCTTGAAGGGGTCTCGGGGGGCGGTGCTGCTGCCCTAGGGGCCACGTC	3880	Db	4901	CTTTGTAGTGTACACGACGGGCTGGTGTGATCTCACACCGCGGAGACACC	4960
Qy	1181	ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle	1200	Qy	1541	ValArgLeuArgAlaTyrPheAsnThrProGlyLeuProValCysGlnAspHisLeuGlu	1560
Db	3881	GTTCGGCTCTTCCGAGCAGCTGTGTCTTCGGGGGCGGTGCTTCCGAGGCGGCAATCCATCGATTTC	3940	Db	4961	GTCCAGCTTAGAGCGTATTTCAACACGCGCGGCTACCGGTGTGTCAAGACCATCTTGA	5020
Qy	1201	ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro	1220	Qy	1561	PheTyrGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThr	1580
Db	3941	CCCGTTCGAGACACTCGAGCTTGTACAGGTCTCCCACTTTCAGTGACACAGCAGCCCA	4000	Db	5021	TTTGTGGAGCGATTTTCAACCGGCTCACACATAGACGCGCCACTTCTCTCCCAACA	5080
Qy	1221	ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLys	1240	Qy	1581	LysGlnSerGlyGluAsnPheAlaTyrLeuThrAlaTyrGlnAlaThrValCysAlaArg	1600
Db	4001	CCGGCTGTGCCCCAGACCTATCAGGTGGGTACTTGCATGCTCCCACTGGCAGTGGAAAG	4060	Db	5081	AACCAAGCGGGGAGAACTTCGGCTACCTAGTAGCTACCAAGCTACGCTGCGCCAGA	5140
Qy	1241	SerThrLysValProValAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnPro	1260	Qy	1601	AlaLysAlaProProProSerTyrAspValMetTyrLysCysLeuThrArgLeuLysPro	1620
Db	4061	AGCACCAAGGTCCTGTGCGGTATGCGCGCCAGGGGTACAAAGTACTAGTGTCTTAACCCC	4120	Db	5141	GCCCAAGGCCCTCCCGCTCTGGGACGCGCATGTGGAAGTGCCTGGCGCGGACTCAAGCCT	5200
Qy	1261	SerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerLysAlaHisGlyIleAsnPro	1280	Qy	1621	ThrLeuValGlyProThrProLeuLeuTyrArgLeuGlySerValThrAsnGluValThr	1640
Db	4121	TCGGTAGCTGCCACCTCGGGTTTGGGGCTACCTATCCAAGGCACATGGCATCAATCCC	4180	Db	5201	ACGCTTTCGGGGCCCAACCTCTCTCTGTATCCGCTTGGGCCCTATTACCAATAGGTACCC	5260
Qy	1281	AsnIleArgThrGlyValArgThrValThrGlyAlaProIleThrTyrSerThrTyr	1300	Qy	1641	LeuThrHisProValThrLysTyrIleAlaThrCysMetGlnAlaAspLeuGluValMet	1660
Db	4181	AACATTAGGACTGGAGTCAGGACCGGTATGACCGGGGAGGCGCATCATCATATCATAT	4240	Db	5261	CTCACACACCTCGGAGCAAGTACATCGCCACATGATGCAAGCTGACCTTGGAGTCATG	5320
Qy	1301	GlyLysPheLeuAlaAspGlyCysAlaGlyGlyValTyrAspIleIleCysAsp	1320	Qy	1661	ThrSerThrTrpValLeuAlaGlyGlyValLeuAlaAlaValAlaAlaTyrCysLeuAla	1680
Db	4241	GGCAAAATTTCTCGCGGATGGGGCTGCGCTAGCGCGGCTTATGATCATCATATGCGAT	4300	Db	5321	ACCAGACGCTGGGTCTCTAGCTGGAGGAGTCTTGGACGCGCTCGCGCATATTGCTGCGG	5380
Qy	1321	GluCysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGlnAla	1340	Qy	1681	ThrGlyCysValCysIleIleGlyArgLeuHisIleAsnGlnArgAlaValAlaPro	1700
Db	4301	GAATGCCACGCTGTGGATGCTACCTCCATTCTCGGCATCGGAAGCTCTTGTATCAAGCA	4360	Db	5381	ACTGGATGCGTTCATCATCGGCGCTTGCACGTCAACCGAGGCTGCTGTGCGCGG	5440
Qy	1341	GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProProGlySerValThr	1360	Qy	1701	AspLysGluValLeuTyrGluAlaPheAspGluMetGluGluCysAlaSerArgAlaAla	1720
Db	4361	GAGACAGCGGGGTGAGACTAATGTGTGCTGCTAGCGGCAACACCCCGGGTCAAGTACA	4420	Db	5441	GATAGAGGCTCTGTATGAGGCTTTTGTAGATGGAGGAATGCGGCTCTAGGCGGCT	5500
				Qy	1721	LeuIleGluGlyGlnArgIleAlaGluMetLeuLysSerLysIleGlnGlyLeuLeu	1740

```
1121 1121 CACATCGATGTTGTTGATGTCGCGCACCTTCTGCTCTGCTCTCTACGCTGGGGACCTC 1180
QY 281 CysGlyGlyValMetLeuAlaGlnMetPheIleValSerProGlnHisIleThrPhe 300
Db 1181 TGTGGCGGGGTGATGTCGCGGCCGAGGTGTTTCACTGCTCGCGCAGTACCACTGGTTT 1240
QY 301 ValGlnAspCysAsnCysSerIleThrProGlyThrIleThrGlyHisArgMetAlaTrp 320
Db 1241 GTGCAGAAATGCAATTTGCTCCATCTACCTTGGGACCATCATCTGGACACCGCATGGCATGG 1300
QY 321 AspMetMetMetAsnTrpSerProThrAlaThrMetIleLeuAlaTrpAlaMetArgVal 340
Db 1301 GACATGATGATGAATGGTCGCGCCACCGCCACCATGATCTGGCGTACGTGATGCGCGTC 1360
QY 341 ProGluValIleIleAspIleIleSerGlyAlaHisIleThrGlyValMetPheGlyLeuAla 360
Db 1361 CCGAGGTTCATCATGACATCGTTAGCGGGCTCACTGGGGGTGTCATGTTGCGCTTGGCC 1420
QY 361 TyrPheSerMetGlnGlyAlaTrpAlaLysValValValIleLeuLeuAlaGly 380
Db 1421 TACTTCTCTATGACGAGGAGGTGGGGAAGTCAATTGTCATCTTCTGCTGGCGCTGGG 1480
QY 381 ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr 400
Db 1481 GTGGACGCGGGCACCAACCGTTGGAGCGCTGTTGCCAGTTCCACCAACGATGATGCC 1540
QY 401 SerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTrp 420
Db 1541 GCGGTGTTGAGCATGCGCCCTCAGCAGAACTTTCAGCTCATTAACACCAACGCGAGTTGG 1600
QY 421 HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer 440
Db 1601 CACATCAACCGTACTGCTTGAATGCAATGACTCTTGAACACCGGCTTCTCGCGGCC 1660
QY 441 LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg 460
Db 1661 TTGTTCTACACCAACCGCTTTAACTCGTCAGGGTGTCCAGGGCGCTGTCCGCTGCCG 1720
QY 461 SerIleGluAlaPheArgValGlyTrpGlyAlaLeuGlnTyrGluAspAsnValThrAsn 480
Db 1721 AACATCGAGGCTTTCGGGATAGGTGGGGCACCCTACAGTACGAGGATTAATGTCAACAA 1780
QY 481 ProGluAspMetArgProTyrCysTrpHisTyrProProArgGlnCysGlyValValSer 500
Db 1781 CAGAGGATATAGGCGGTACTCTGGCACTACCCCCCAAGACCGGTGTGGGTAGTCCCC 1840
QY 501 AlaLysThrValCysGlyProValTyrCysPheThrProSerProValValGlyThr 520
Db 1841 GCGAGGTCTGTGTGGCCAGTGTTACTGTTTCACCCCGCCAGCCGCTAGTGTGGGCAG 1900
QY 521 ThrAspArgLeuGlyAlaProThrTyrThrTrpGlyGluAsnGluThrAspValPheLeu 540
Db 1901 ACCGACAGAGCTGGAGTGGCCACCTACATGAGGGAGAGAAATGAGACAGATGCTTCTCTA 1960
QY 541 LeuAsnSerThrArgProLeuGlySerTrpPheGlyCysThrTrpMetAsnSerSer 560
Db 1961 CTGAACAGCACCCGACCGCGAGGCTCATGTTTCCGCTGACGCGGATGAATCACT 2020
QY 561 GlyTyrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer 580
Db 2021 GGTTCACCAAGACTTGTGGCGGCCACCTTCCGCGCCAGAGCTGACTTCAACGCCAGC 2080
QY 581 ThrAspLeuLeuCysProThrAspCysPheArgLysHisProAspThrThrTyrLeuLys 600
Db 2081 ACGGACTTGTGTGGCTACGAGTCTTTTAGAAGCATCTCGATGCCACTTATATAAG 2140
QY 601 CysGlySerGlyProThrLeuThrProArgCysLeuIleAspTyrProTyrArgLeuTrp 620
Db 2141 TGTGGTCTTGTGGGCCCTGGCTCACCAAAAGTGGTCCACTTACCTTACAGACTCTGG 2200
QY 621 HisTyrProCysThrValAsnTyrThrIlePheLysIleArgMetTyrValGlyGlyVal 640
```

Claim 5; Page 13-25; 36pp; Japanese.

The invention relates to a human polypeptide related to hepatitis C virus (HCV), and the polynucleotide encoding it. The polypeptide can be used for the development of gene therapy on fulminant hepatitis C. This sequence represents cDNA encoding a human HCV-related polypeptide.

Sequence 9678 BP: 1948 A; 2915 C; 2723 G; 2092 T; 0 other;

Alignment Scores:			
Pred.	No. :		
0	14943.00	Length:	9678
Score:	95.42%	Matches:	2765
Percent Similarity:	Conservative:		129
Best Local Similarity:	Mismatches:		139
91.16%	Indels:		0
Query Match:	Gaps:		0
24	DB:		
24			

US-09-980-559-2 (1-3033) x ABK88904 (1-9678)

Qy	1	MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgProGln	20
Db	341	ATGAGCAAAATCCTAAACCTCAAAAGAAAAACCAAAAGAAACACCAACCGTCCGCCAGAA	400
Qy	21	AspValLysPheProGlyGlyGlyGlnIleValGlyGlyValTyLeuLeuProArgArg	40
Db	401	GACGTTAAAGTTCCCGGGCGGGCGGCAGATCGTTGGCGGAGTATACTTGTTCGGCGCAGG	460
Qy	41	GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly	60
Db	461	GGCCCCAGGTTGGGTGTGGCAGCAGACAGGAAAACTTCGGAGCGGTCCCGACCACTGGG	520
Qy	61	ArgArgGlnProIleProLysAspArgArgSerThrGlyLysSerTrpGlyLysProGly	80
Db	521	AGACGCCAGCCATCCCCAAAGATCGGGCTCCACTGGCAAGCCCTGGGGGAAAAACAGGT	580
Qy	81	TyrProTrpProLeuTyrGlyAsnGluGlyLeuGlyTrpAlaGlyTrpLeuLeuSerPro	100
Db	581	CGCCCTCGCCCTATATGGGAATGAGGAGCTCGGCTGGCGAGGATGGTCTCTGTCCCC	640
Qy	101	ArgGlySerArgProSerTrpGlyProAsnAspProArgHisArgSerArgAsnValGly	120
Db	641	CGAGGCTCTCGCCCTCTCTGGGGCCCCACTGACCCCGCGCATAGTTCGCGCAACGTGGGT	700
Qy	121	LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyTrileProValVal	140
Db	701	AAAGTCATCGACACCCCTAACGTGTGGCTTTGCCGACCTCATGGGGTACATCCCCGTCGTA	760
Qy	141	GlyAlaProLeuGlyGlyValAlaArgAlaLeuAlaHisGlyValArgValLeuGluAsp	160
Db	761	GGCGCCCGCTTAGTGGCGCCGCCAGAGCTGTCCGCGCAGCGGTGAGAGTCTCTGGAGGAC	820
Qy	161	GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla	180
Db	821	GGGGTTAATTATGCAACAGGGAACCTTACC CGGTTCCTCCCTTTCTATCTTCTTGTCTGGCC	880
Qy	181	LeuLeuSerCysIleThrProValSerAlaAlaGluValLysAsnIleSerThrGly	200
Db	881	CTGTTGTCTCTGCATCACCGTTCCGGTCTCTGTGCCAGGTGAGAGTACCAGTAGCAGC	940
Qy	201	TyrMetValThrAsnAspCysThrAsnAspSerIleThrTrpGlnLeuGlnAlaVal	220
Db	941	TACATGGTGACCAATGACTGCTCCAATGACAGCATCACTTGGCAGCTCGAGGTCGGGTT	1000
Qy	221	LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTrpIle	240
Db	1001	CTCCAGTCCCGGGTGGCGTCCGTCGAGAGAGTGGGAAATACGTACCGGTCTGGGTG	1060
Qy	241	ProValSerProAsnValAlaValGlnArgProGlyAlaLeuThrGlnGlyLeuArgThr	260
Db	1061	CAGTCTCCGCCAAACATGGCTGTGGCGGACGCCCGGTGCCCTCACGCGAGGCTCGCGGACG	1120
Qy	261	HisIleAspMetValValMetSerAlaThrLeuCysSerAlaLeuTyrValGlyAspLeu	280

RESULT 5

ABK88904
ID ABK88904 standard: cDNA: 9678 BP.

AC ABK88904:

07-OCT-2002 (first entry)

DE Human HCV-related polypeptide encoding cDNA.

Hepatitis C virus: human: ss: virucide: gene therapy: HCV: KW

fulminant hepatitis C.

XX

XX

FH	Key	Location/Comments
000		000
001		001
002		002
003		003
004		004
005		005
006		006
007		007
008		008
009		009
010		010
011		011
012		012
013		013
014		014
015		015
016		016
017		017
018		018
019		019
020		020
021		021
022		022
023		023
024		024
025		025
026		026
027		027
028		028
029		029
030		030
031		031
032		032
033		033
034		034
035		035
036		036
037		037
038		038
039		039
040		040
041		041
042		042
043		043
044		044
045		045
046		046
047		047
048		048
049		049
050		050
051		051
052		052
053		053
054		054
055		055
056		056
057		057
058		058
059		059
060		060
061		061
062		062
063		063
064		064
065		065
066		066
067		067
068		068
069		069
070		070
071		071
072		072
073		073
074		074
075		075
076		076
077		077
078		078
079		079
080		080
081		081
082		082
083		083
084		084
085		085
086		086
087		087
088		088
089		089
090		090
091		091
092		092
093		093
094		094
095		095
096		096
097		097
098		098
099		099

FT CDS 341..9442 / *tag= a

```

FT      /product= "Human HCV-related protein"
FT      /cay- a

```

JP2002171978-A.

18-JUN-2002:

01-DEC-2000; 2000JP-0367365.

01-DEC-2000: 2000JP-0367365.

XX PA (TOKR-) ZH TOKYOTO RINSHO IGAKI SOGO KENKYUISHO.

PA (TORA) TORAY IND

WPI; 2002-569884/61.

DR P-PSDB; ABG30688.

XX

PT A gene of a fulminant hepatitis C virus strain and the encoded

PI polypeptide useful in gene therapy to treat hepatitis C -

Copied from 09540843 on 05/19/2004

Db 6701 CCTTCTCCAGAGTCTTTCTCGGTGGACGGAGTGCAGATCCATAGGTTTGCCCCCATTA 6760
Qy ProLysProPheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal 2160
Db 6761 CCGAGCCGTTTTTGGGATGAGGCTCGTCTGGGTGGGCTTAATTCATTGTTCGTC 6820
Qy GlySerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThr 2180
Db 6821 GGGTCTCAGCTCCCTTGGCATCTCGAACCTGCACAGACGCTATTGACGTCCTAACA 6880
Qy AspProSerHisIleThrAlaGluThrAlaAlaArgArgLeuAlaArgGlySerProPro 2200
Db 6881 GACCCATCCCATATACGGCGAGACTGCAGCGGGGCTTGGCAGCGGGGTCCACCCCG 6940
Qy SerGluAlaSerSerAlaSerGlnLeuSerAlaProSerLeuAlaThrCysThr 2220
Db 6941 TCCAGGCAAGCTCTCAGCGAGCCAGCTATCGGCACCATCGCTGGAGCCACTGCACC 7000
Qy ThrHisGlyLysAlaIleAspValAspMetValAspAlaAsnLeuPheMetGlyGlyAsp 2240
Db 7001 ACCCAGCGCAAGCCATATGATGGACATGGTGGATGCCAACCTGTTTCATGGGGGGGAT 7060
Qy ValThrArgIleGluSerGlySerLysValValValLeuAspSerLeuAspProMetVal 2260
Db 7061 GTGACCCGGATAGAGTCTGAGTCCAAAGTGGTGGTCTCTGGACTCTCTCGACCCCAATGGTC 7120
Qy GluGluArgSerAspLeuGluProSerIleProSerGluTyrMetLeuProLysLysArg 2280
Db 7121 GAAGAAGGAGGACCTTGAGCTTCGATACCATCGGAATATATGCTCCCAAGAAGAGA 7180
Qy PheProProAlaLeuProAlaIleAspValAspTyrAsnProProLeuValGluSer 2300
Db 7181 TTCCACCAACCTTACCGCTTGGGCAAGCCGCTGATTACACCCACCGCTTGTGGATCG 7240
Qy TrpLysArgProAspTyrGlnProAlaThrValAlaGlyCysAlaLeuProProProArg 2320
Db 7241 TGGAAAGGCGAGATTACCAACCGGCGACCTGTTGGCGGCTCGCTCTCCCGCCCGCTAAG 7300
Qy LysThrProThrProProProArgArgArgArgThrValGlyLeuSerGluAspSerIle 2340
Db 7301 AAAACCCGAGCGCTCCCGCAAGGAGCGCGGACAGTGGGTCTGAGTGAGAGCTCCATA 7360
Qy GlyAspAlaLeuGlnGlnLeuAlaIleLysSerPheGlyGlnProProProSerGlyAsp 2360
Db 7361 GCAGATGCCCTACACAGCTGGCCATCAAGTCTTTGGCCAGCGCCCGCCCGCGAT 7420
Qy SerGlyLeuSerThrGlyAlaGlyAlaAlaAspSerGlySerGlnThrProProAspGlu 2380
Db 7421 TCAGGCGCTTTCACGGGGCGGACGCGACCGGATTCGGCAGTCCGACGCGCCCGCATGAG 7480
Qy LeuAlaLeuSerGluThrGlySerIleSerSerMetProProLeuGluGlyLeuGly 2400
Db 7481 TTGGCCCTTTTCGAGACAGGTTCCATCTCTCATGCCCGCTCTCGAGGGGAGCGCTGGA 7540
Qy AspProAspLeuGluProGluGlnValGluProGlnProProGlnGlyValAla 2420
Db 7541 GATCCAGACTTGAGGCTTGAGGAGGTAGAGCTTCAACCTTCCCCCGGGGGGGTGGTA 7600
Qy AlaProGlySerAspSerGlySerTrpSerThrCysSerGluGluAspAspSerValVal 2440
Db 7601 ACCCGCGGCTCAGGCTCGGGGTCTTGGTCTTACTTGTCTCGAGGAGGACGACTCCGCTG 7660
Qy CysCysSerMetSerTrpThrThrGlyAlaLeuIleThrProCysSerProGluGlu 2460
Db 7661 TGCTGCTCATGTCATCTCCGACCGGGGCTCTAATACTCTTGTAGCCCGGAGAG 7720
Qy GluLysLeuProIleAsnProLeuSerAsnSerLeuLeuArgTyrHisLeuLysValTyr 2480
Db 7721 GAAAGTTGCCAATTGGCCCTTGGCACTCCCTGTTGGGATATCACACACAGGTGTAC 7780
Qy CysThrThrThrLysSerAlaSerLeuArgAlaLysLysValThrPheAspArgMetGln 2500
Db

Db 7781 TGTACCACATCAAGAGCGCCCTCATTAAGGGCTAAAGAGTAACCTTTTGTATAGATGCAA 7840
Qy ValLeuAspSerTyrTyrAspSerValLeuLysAspIleLysLeuAlaLalaSerLysVal 2520
Db 7841 GGGCTCGAGCGCTCATTTATGACTCAGCTTTGAAGGACATTAAGCTAGCGGCTTCAAGGTC 7900
Qy ThrAlaArgLeuLeuThrMetGluGluAlaCysGlnLeuThrProProHisSerAlaArg 2540
Db 7901 ACCGCAAGGCTTCTCACTTTAGAGGAGGCTGCGAGTTAACTCCACCCACCTCTGCAAGA 7960
Qy SerLysTyrGlyPheClyAlaLysGluValArgSerLeuSerClyArgAlaValAsnHis 2560
Db 7961 TCCAAATATGGGTTTGGGGCTTAAGGAGGTCGCGAGCTTGTCCGGGAGAGCCGTTAACCCAC 8020
Qy IleLysSerValTrpLysAspLeuLeuGluAspSerGluThrProIleProThrThrIle 2580
Db 8021 ATCAAGTCCGTGTGGAGGACCTCTCTGGAAGACACACAAACACCAATCTCTACACCAATC 8080
Qy MetAlaLysAsnGluValPheCysValAspProThrLysGlyLysLysAlaLalaArg 2600
Db 8081 ATGGCCAAAATGAGGTGTTCTCGGTGGACCCACCAAGGGGGTAAAGAACAGCTCGC 8140
Qy LeuIleValTyrProAspLeuGlyValArgValCysGluLysMetAlaLeuTyrAspIle 2620
Db 8141 CTTATCGTTTACCTCGACCTCGGCGTCAAGGCTGCGAGAAATGGCCCTTTATGATATC 8200
Qy ThrGlnLysLeuProGluAlaValMetGlyAlaSerTyrGlyPheGlnTyrSerProAla 2640
Db 8201 ACACAAAAGCTTCTCAGCGGTGATGGGGCTTCTTATGGATTTCCAGTACTCCCGCT 8260
Qy GlnArgValGluPheLeuLeuLysAlaTrpAlaGluLysLysAspProMetGlyPheSer 2660
Db 8261 CAGCGGTGAGTCTTCTTTGAAGCATGCGCGGAAAGAAAGACCTATGGGTTTTTCG 8320
Qy TyrAspThrArgCysPheAspSerThrValThrGluArgAspIleArgThrGluGluSer 2680
Db 8321 TATGATACCCGATGCTTTGACTCAACCGCTCACTGAGAGAGACATCAGAGCTGAGGAGTCC 8380
Qy IleTyrArgAlaCysSerLeuProGluAlaHisThrAlaIleHisSerLeuThrGlu 2700
Db 8381 ATATATCGGGCTTGTTCCTTCCGAGGAGGCGCACCTGCCATACACTCACTGACTGAG 8440
Qy ArgLeuTyrValGlyProMetPheAsnSerLysGlyGlnThrCysGlyTyrArgArg 2720
Db 8441 AGACTTACGTGGAGGGGCCCATGTTCAACACAGGGCCAGACCTCGCGGTACAGCGT 8500
Qy CysArgAlaSerGlyValLeuThrThrSerMetGlyAsnThrIleThrCysTyrValLys 2740
Db 8501 TGCCGCGCCAGCGGGGTGCTTACCACCTAGCATGGGGAACACCATCACATGCTATGTGAAA 8560
Qy AlaLeuAlaAlaCysLysAlaAlaGlyIleIleAlaProThrMetLeuValCysGlyAsp 2760
Db 8561 GCTTGTAGCGGCTGTAAAGGCTCGAGGATAATTGGCCGCCCAATGCTGTGTGCGCGAT 8620
Qy AspLeuValValIleSerGluSerGlnGlyThrGluGluAspGluArgAsnLeuArgAla 2780
Db 8621 GACTGTTGTTCATCTCAGAGAGCCAGGGGACCGAGGAGGACGCGGGAACCTGAGAGCC 8680
Qy PheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProArgProGluTyr 2800
Db 8681 TTCAGGAGGCTATCACCAGGTATTTCTGCCCTCTCTGTTGAGTACCCCGCCAGCGGAATAT 8740
Qy AspLeuGluLeuIleThrSerCysSerAsnValSerValAlaLeuGlyProGlnGly 2820
Db 8741 GACCTGGAGCTGATTAACATTTGCTTCTCAATGTCTGTGGCGTGTGGGCCACCAAGGC 8800
Qy ArgArgArgTyrTyrLeuThrArgAspProThrThrProIleAlaArgAlaIleTrpGlu 2840
Db 8801 CGCCGAGACTACTACTGACGAGACCTTACCATCCCAATCGCCGGGCTGCGCTGGGA 8860
Qy ThrValArgHisSerProValAsnSerTrpLeuGlyAsnIleIleGlnTyrAlaProThr 2860
Db 8861 ACAGTTAGACACTCCCGCTGTCAATTCTATGGCTTAGGAAACATCATCCAGTACGCCCCAACCC 8920

Qy	1401	ValLeuValSerCysAspGluLeuAlaAlaLeuArgGlyMetGlyLeuAsnSerValAla	1420	Qy	1761	ValGluGlnPheTrpAlaLysHisMetTrpAsnPhelleSerGlyIleGlnTyrLeuAla	1780
Db	4541	AGAAAAGTGTGACAGCTCGCGGGCCCTTCGGGTATGGCTTGAACGAGTGGCA	4600	Db	5621	GTAGAGCAATTCGTGGCCAAACACATGTGAACTTCATCAGCGCATTCATACACCTCGCA	5680
Qy	1421	TyrTyrArgGlyLeuAspValSerValIleProThrGlnGlyAspValValValAla	1440	Qy	1781	GlyLeuSerThrLeuProGlyAsnProAlaValAlaSerMetMetAlaPheSerAlaAla	1800
Db	4601	TACTACAGAGGGCTGGACGTCTCCGTAATACCAACTCAGGGACGCTAGTGTGTCGCC	4660	Db	5681	GGCATATCAACACTGCGAGGAACCTGCTGTAGCTTCCATGATGCAATTCAGTGGCGCC	5740
Qy	1441	ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnVal	1460	Qy	1801	LeuThrSerProLeuSerThrThrIleLeuLeuAsnIleLeuGlyGlyTyrLeu	1820
Db	4661	ACCGAGCCCTCATGACGGGTTTACTGGAGACTTTGACTCCGTGATCGACTCAACGTA	4720	Db	5741	CTCACAGCTCCGTTGTCACTAGCACCATCTCTCTCAACATTTTGGGGGCTGGCTA	5800
Qy	1461	AlaValThrGlnValValAspPheSerLeuAspProThrPheThrIleThrGlnIle	1480	Qy	1821	AlaSerGlnIleAlaProProAlaGlyAlaThrGlyPheValValSerGlyLeuValGly	1840
Db	4721	GCCTCTACTCAAGTTGTAGACTTCAGCTTGACCCCCCATTCACATACCAACACAGACT	4780	Db	5801	GCATCCCAATTTGGCTCCCGGGGCTACCGGCTTCTGTCAGTGGCTGGTGGG	5860
Qy	1481	ValProGlnAspAlaValSerArgSerGlnArgGlyThrGlyArgGlyArgLeu	1500	Qy	1841	AlaAlaValGlySerIleGlyLeuGlyLysValLeuValAspIleLeuAlaGlyTyrGly	1860
Db	4781	GTCCCTCAACAGCGCTCTCACGTAGCCAGCGCGGGCCGACGGGACGAGACTG	4840	Db	5861	GCTGCCGTAGGAGCATAGGCTTGGGTAAAGGTGCTGGATCCCTGGCAGGGTATGGT	5920
Qy	1501	GlyIleTyrArgTyrValSerThrGlyGluArgAlaSerGlyMetPheAspSerValVal	1520	Qy	1861	AlaGlyIleSerGlyAlaLeuValAlaPheLysIleMetSerGlyGlyLysProSerMet	1880
Db	4841	GGTATTTATAGGTATGTTTCCACTGGTGTAGCGAGCCCTCAGCAATGTTTGTACGTAGTG	4900	Db	5921	GGGGCATTTTCGGGGCTCTCTGTCGATTCAGATCATCTCTGGCGAAGAGCCCTCCATG	5980
Qy	1521	LeuCysGluCysTyrAspAlaGlyAlaAlaTyrTyrGluLeuThrProSerGluThrThr	1540	Qy	1881	GluAspValValAsnLeuLeuProGlyIleLeuSerProGlyAlaLeuValValGlyVal	1900
Db	4901	CTCTGCCAGTGTACGATGACGAGGCGCGATGATGAGTTCACACGCGGAGACCACT	4960	Db	5981	GAGGATGTTGTCAACCTGCTGCTGGAAATCTCTCTCCGGTCCCTGGTGGTGGAGTC	6040
Qy	1541	ValArgLeuArgAlaTyrPheAsnThrProGlyLeuProValCysGlnAspHisLeuGlu	1560	Qy	1901	IleCysAlaAlaIleLeuArgArgHisValGlyProGlyGlyGlyAlaValGlnTyrMet	1920
Db	4961	GTCCAGCTCAGAGCATATTTCAACACACCTGTTTGGCTGTGTGCCAAGACCATCTTGAG	5020	Db	6041	ATCTGGCGCGCCATCTCTGCGGACACGCTGGGACCGGGGAGGCGCTGTCCAAATGGATG	6100
Qy	1561	PheTrpGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThr	1580	Qy	1921	AsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaAlaProThrHisTyrValThr	1940
Db	5021	TTTTGGAGGAGCAGTTTTTACCGGCTCCACACACATAGATCCCACTTCTTCCCAACA	5080	Db	6101	AATAGCTCATTTGCTTTCAGAGGAAACACGCTGCCGCCACCCACCTACGTGACG	6160
Qy	1581	LysGlnSerGlyGluAsnPheAlaTyrLeuThrAlaTyrGlnAlaThrValCysAlaArg	1600	Qy	1941	GluSerAspAlaSerGlnArgValThrGlnLeuLeuGlySerLeuThrIleThrSerLeu	1960
Db	5081	AGCAATCGGGGGAATTTTCCGATATTTAACAGCTTACAGGCTACAGTGTGGCTAGG	5140	Db	6161	GAGTCGGATGCGTCGAGCGGTGTGACCAACTACTTGGCTCCCTTACCATAACAGCCTG	6220
Qy	1601	AlaLysAlaProProSerTrpAspValMetTrpLysCysLeuThrArgLeuLysPro	1620	Qy	1961	LeuArgArgLeuHisAsnTrpIleThrGluAspCysProIleProCysGlySerTyr	1980
Db	5141	GCCAAAGCCCCCCCCCTCTGGAGCGTCTATGTGGAAGTGTGTGACTCAAGCCC	5200	Db	6221	CTCAGAGACTCCACAACTGGATTTACTGAAGACTGCCCACTCCATGCGAGCGCTCTGG	6280
Qy	1621	ThrLeuValGlyProThrProLeuLeuTyrArgLeuGlySerValThrAsnGluValThr	1640	Qy	1981	LeuArgAspValTrpAspTrpValCysThrIleLeuThrAspPheLysAsnTrpLeuThr	2000
Db	5201	ACACTCGTGGGCCCCACACCTCTCTGTACCGCTTGGGCTCTGTATACCAAGAGTCACT	5260	Db	6281	CTCCGCGATGTGGGATGGGTTTGACCATCTTAACAGACTTTAAAAAATGGTGACC	6340
Qy	1641	LeuThrHisProValThrLysTyrIleAlaThrCysMetGlnAlaAspLeuGluValMet	1660	Qy	2001	SerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTyrLys	2020
Db	5261	CTACGATCTCTGTGAGAAATACATCGCACTTGCATGCAAGCCGACCTTGAGTCTATG	5320	Db	6341	TCCAAATTTGTTCCAAAGATGCTGCTCTCCCTTTTATCTCTTGTCAAAAGGGGTACAAG	6400
Qy	1661	ThrSerThrTrpValLeuAlaGlyGlyValLeuAlaAlaValAlaAlaTyrCysLeuAla	1680	Qy	2021	GlyValTrpAlaGlyThrGlyIleMetThrThrArgCysProCysGlyAlaAsnIleSer	2040
Db	5321	ACCAAGCAGTGGGTCTTAGCTGGGGGGTCTTGGCGCGCTGCGCGGTACTGCTGCG	5380	Db	6401	GGCGTGTGGGCTGGCATCTGGTATCATGACACACGCTGCTTGGCGCGCAATATCTCT	6460
Qy	1681	ThrGlyCysValCysIleIleGlyArgLeuHisIleAsnGlnArgAlaValValAlaPro	1700	Qy	2041	GlyAsnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTyr	2060
Db	5381	ACCGGGTGTGTTCATCATCGCGCTTGCACGTAAACCAAGCAGCGCGCTGTTGCACCG	5440	Db	6461	GGCAATGTCGCTGGCTCCATGAGAAATTTACGGGGCCCAAAACCTTGATGAATATCTCG	6520
Qy	1701	AspLysGluValLeuTyrGluAlaPheAspGluMetGluGluCysAlaSerArgAlaAla	1720	Qy	2061	GlnGlyThrPheProIleAsnCysTyrThrGluGlyGlnCysValProLysProAlaPro	2080
Db	5441	GACAAGGAGGTCCTTATAGAGCTTTTGTATGAGATGAGGAATGTGCTCTTAGAGCGCT	5500	Db	6521	CAGGGGACCTTTCCCATCAATTTGTTACACGGGGCCAGTGGCTGCCGAAACCCGACCA	6580
Qy	1721	LeuIleGluGluGlyGlnArgIleAlaGluMetLeuLysSerLysIleGlnGlyLeu	1740	Qy	2081	AsnPheLysValAlaIleTrpArgValAlaAlaSerGluTyrAlaGluValThrGlnHis	2100
Db	5501	CTCATTAAGAGGGGACGCGATAGCGAGATGTGAAGTCCAAAGATCCAAAGGCTTATTG	5560	Db	6581	AACTTTAAGATCCCATCTGGAGGGTGGCGCTCCAGAGTACGCGGAGGTGACGACAC	6640
Qy	1741	GlnGlnAlaSerLysGlnAlaGlnAspIleGlnProThrValGlnAlaSerTrpProLys	1760	Qy	2101	GlySerTyrHisTyrIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeu	2120
Db	5561	CAGCAAGCTTCCAAACAGCTCAAGACATACACCGCTGTGAGGCTTCTTGGCCCAAG	5620	Db	6641	GGGTATACCACTTACATACAGGACTTACCACTGTATAACTTTGAAAGTTCCTTGGCACTA	6700
				Qy	2121	ProSerProGluPhePheSerTrpValAspGlyValGlnIleHisArgPheAlaProThr	2140

Db	2321	AGAGACAGAGTCAACTGCTCTCTTGGTGCACCTCCACACGAGTGGGCCATTTTACCT	2380	Db	3401	GGCTATACCTCCAAAGGGTGGAGTCTTCTCGCCCCCATCACTGCTTATGCCACGACACA	3460
Qy	681	CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisLeuHisGlnAniLe	700	Qy	1041	ArgGlyLeuLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGluGlnAla	1060
Db	2381	TGCATTTACTCGGACCTGCCCGCTTGTGCACCTGGTCTTCCACCTCCACCAAAACATC	2440	Db	3461	CGGGCCCTTTGGGACCATAGTGGTGGACATGACGGGGCGGCAAGAACAGAACAGGCC	3520
Qy	701	ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrLysTyrIleValArgTyr	720	Qy	1061	GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly	1080
Db	2441	GTGACGTCGCAATTCATGATGCGCTATCCTGCTCTCACAATAATACATCGTCCGATGG	2500	Db	3521	GGGGAGATTTCAGTCTCTGTCACCGTCACTCAGTCTCTCTCGGAACAACCATCTCGGGG	3580
Qy	721	GluTrpValIleLeuLeuPheLeuLeuLeuAlaAspAlaArgValCysAlaCysLeuTrp	740	Qy	1081	ValLeuTrpThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro	1100
Db	2501	GAGTGGGTAGTACTTATTCCTGCTCTTAGCGGACGCCAGGCTTGGCGCTCTGCTATGG	2560	Db	3581	GTCTATTGACTGTCTACCATGGAGCTGGCAACAAGACTCTAGCCGGCTCAGCGGTCGG	3640
Qy	741	MetLeuIleLeuLeuGlyGlnAlaGluAlaAlaLeuGluLysLeuValIleLeuHisAla	760	Qy	1101	ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTrpProSerProGlyGly	1120
Db	2561	ATGCTCATCTGTGTGGCCAGCGGAGAGAGAGCTAGAGAAAGTGTGCTGTTCGACGCT	2620	Db	3641	GTCCACACAGATGCTCTCCAGTCTGAGGGGGAGCTTAGTGGGGTGGCCAGCCCCCGGG	3700
Qy	761	AlaSerAlaAlaSerCysAsnGlyPheLeuTyrPheValIlePhePheValAlaAlaTrp	780	Qy	1121	ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArgAsn	1140
Db	2621	GCGAGCGCAGTAGCTCAATGCTTCTTATATCTTGTATCTTTTGTGGCTGCTTGG	2680	Db	3701	ACCAATCTTTGAGCGCTGCACGCTGGAGCGGTGACCTATACCTGGTCACGCGAAC	3760
Qy	781	TyrIleLysGlyArgValValProLeuAlaThrTyrSerLeuThrGlyLeuTrpSerPhe	800	Qy	1141	AlaAspValIleProAlaArgArgGlyAspLysArgGlyAlaLeuLeuSerProArg	1160
Db	2681	TACATCAAGGGTCGGGTAGTCCCTTGGCTACTTATTCCTCACTGGCCTATGGTCTTT	2740	Db	3761	GCTGATGTCTATCCCGCTCGAAGACGCGGGGACAAGCGAGGAGCGCTACTCTCCCGAGA	3820
Qy	801	SerLeuLeuLeuAlaLeuProGlnGlnAlaTyrAlaTyrAspAlaSerValHisGly	820	Qy	1161	ProLeuSerThrLeuLysGlySerSerGlyGlyProValLeuCysProArgGlyHisAla	1180
Db	2741	GGCTACTGCTCTAGCATGGCCCCAACAGGCTTATGCTTATGACGCATCTGTACATGGT	2800	Db	3821	CCTCTTTCCACCTTGAAGGGGTCTCGGGGGGGCCGGTCTCTGCCCCAGAGGCCACGT	3880
Qy	821	GlnIleGlyAlaAlaLeuLeuValMetIleThrLeuPheThrLeuThrProGlyTyrLys	840	Qy	1181	ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle	1200
Db	2801	CAGATAGGACGACTCTGTGGTGTACTGATGATCTCTTTTACACTCACCCCGGGATTAAG	2860	Db	3881	GTCCGGGTCTTCCGGGACGCGCTGTGCTCCCGGGGGTGGCCAAAGTCTCATATTTATC	3940
Qy	841	ThrLeuLeuSerArgPheLeuTrpTrpLeuCysTyrLeuLeuThrLeuGlyGluAlaMet	860	Qy	1201	ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro	1220
Db	2861	ACCTTCTCAGCCGGTCTGTGGTGTGTGCTATCTTCTGACCCCTGGCGGAAGCTATG	2920	Db	3941	CCCGTTGAGACACTTGACATCGTCACTCGTCCCCCACCTTTAGTGACAACAGCACCA	4000
Qy	861	ValGlnGluTrpAlaProMetGlnValArgGlyGlyArgAspGlyIleIleTrpAla	880	Qy	1221	ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLys	1240
Db	2921	GTCCAGGAGTGGGCACACCTATGACAGGTGCGCGGTGGCGGTGATGATATATGGGC	2980	Db	4001	CCTGCTGTGCCCCAACTTATCAGTGGGTACTTACATGCCCCGACGCTGCTAGTGAAG	4060
Qy	881	ValAlaIlePheTyrProGlyValValPheAspIleThrLysTrpLeuLeuAlaValLeu	900	Qy	1241	SerThrLysValProValAlaTyrAlaGlnGlyTyrLysValLeuValLeuAsnPro	1260
Db	2981	GTCCCATATTCGCGGGGTGGTGTGTGATTAACCAAGTGGTCTTGGCGGTCTT	3040	Db	4061	AGCACCAAGTCTCTGTGCGGTATGCGCTCAGGGGTACAAAGTGTCTAGTGTATATCC	4120
Qy	901	GlyProAlaTyrLeuLeuLysGlyAlaLeuThrArgValProTyrPheValArgAlaHis	920	Qy	1261	SerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerLysAlaHisGlyIleAsnPro	1280
Db	3041	GGGCTTGTCTATCTCTAAAGGTGTCTTGAAGGTGTGCGGTGTCTGTCAGGGCTCAC	3100	Db	4121	TCGGTGGCTGCCACCTCGGGTTCGGGGTTCGGGGTTCGGGGTTCGGGGTTCGGGGT	4180
Qy	921	AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyGlyArgTyrValGlnMet	940	Qy	1281	AniIleArgThrGlyValArgThrValThrThrGlyAlaProIleThrTyrSerThrTyr	1300
Db	3101	GCTCTACTAAGGATGTGCACCATGGTAAGGATCTCGCGGGGGTAGGTACGTTCAGATG	3160	Db	4181	AACATTAGACTTGGGGTTCAGGACTGTGACCGGGGGGCCCATCAGCTACTCCATAT	4240
Qy	941	AlaLeuLeuAlaLeuGlyArgTyrThrGlyThrTyrIleTyrAspHisLeuThrProMet	960	Qy	1301	GlyLysPheLeuAlaAspGlyGlyCysAlaGlyGlyAlaTyrAspIleIleCysAsp	1320
Db	3161	GTGCTACTAGCTTGGCAGTGGACTGACCTTATCATGACCACTTACCCCTATG	3220	Db	4241	GGCAATTCCTCGCGATGGGGCTGCGCAGCGGGGGCTATGACATCATATATGCGAT	4300
Qy	961	SerAspTrpAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIleIlePhe	980	Qy	1321	GluCysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGlnAla	1340
Db	3221	TCGATTGGCTGCTAATGGCTTGGCGGCTTGGCGGTGGCGGTGGCGGTGGCGGTGGCG	3280	Db	4301	GAATGCCATCGCGTGGACTTACCACCATTCCTCGGCATCGGAACAGTCTCGATCAAGCA	4360
Qy	981	SerProMetGluLysLysValIleValTrpGlyAlaGluThrAlaAlaCysGlyAspIle	1000	Qy	1341	GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProGlySerValThr	1360
Db	3281	AGTCCGATGGAGAAAAAGTCACTGCTGGGGAGCGGAGACAGCTGCTTGGCGGGATATC	3340	Db	4361	GAGACAGCGGGGTTCAGGCTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4420
Qy	1001	LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp	1020	Qy	1361	ThrProHisProAsnIleGluGluValAlaLeuGlyGlnGluGlyGluIleProPheTyr	1380
Db	3341	TTACACGGACTTCCCGTGGCGGCGACTTGGCGGGAGGCTCTCTCTGCGCCAGCTGAT	3400	Db	4421	ACCCCCACCCCAACATAGAGAGGTGGCTTCCGGCAGGAGGGTGGAGATCCCTCTTAT	4480
Qy	1021	GlyTyrThrSerLysGlyTrpSerLeuLeuAlaProIleThrAlaTyrAlaGlnThr	1040	Qy	1381	GlyArgAlaIleProLeuSerTyrIleLysGlyGlyArgHisLeuIlePheCysHisSer	1400
Db				Db	4481	GGGAGGGGATTCCTCTGTCATACATCAAGGGAGGAGACATTTGATCTTCTGCTCACTCA	4540

XX Sequence 9589 BP; 1968 A; 2820 C; 2635 G; 2166 T; 0 other;

Alignment Scores:

Pred. No.: 0 Length: 9589
 Score: 15824.00 Matches: 2966
 Percent Similarity: 98.78% Conservative: 30
 Best Local Similarity: 97.79% Mismatches: 37
 Query Match: 98.16% Indels: 0
 DB: 14 Gaps: 0

US-09-980-559-2 (1-3033) x AAQ38218 (1-9589)

QY 1 MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGln 20
 DB 341 ATGAGCACAATCTTAACCTCAAGAAACCAAGAAACCAACCAACCTGCCCCACAA 400
 QY 21 AspValLysPheProGlyGlyGlnValGlyValGlyValGlyValGlyValGlyValGly 40
 DB 401 GACGTTAAAGTTCCGGCGCGCGCAGATCGTTGGCGAGTATATCTTTGGCGCGCAGG 460
 QY 41 GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly 60
 DB 461 GGGCCCCAGGTTGGTGTGCGCGCACAAAGGAAGACTTCGGAGCGGTCCCGACCGTGA 520
 QY 61 ArgArgGlnProLysPheProGlyGlyGlnValGlyValGlyValGlyValGlyValGly 80
 DB 521 AGCGCCAGCCCATCCCTAAGGATCGCGCTCACTGGCAATCTGGGGAACCAAGGA 580
 QY 81 TyrProTrpProLeuTyrGlyAsnGluGlyLeuGlyTrpAlaGlyTrpLeuLeuSerPro 100
 DB 581 TACCCCTGGCCCTATACGGGAATGAGGACTCGGCTGGCGAGATGGCTCTGTGCCCC 640
 QY 101 ArgGlySerArgProSerTrpGlyProAsnAspProArgHisArgSerArgAsnValGly 120
 DB 641 CGAGGTTCCCGTCCCTCTGGGCGCCCAATGACCCCGCATAGGTCCTCCGCAACGTTGG 700
 QY 121 LysValLysLeuThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrLysProValVal 140
 DB 701 AAGGTCAATGATACCTTAAGCGTGGGCTTTGGCCGACCTCAATGGGTACATCCCTGTG 760
 QY 141 GlyAlaProLeuGlyGlyValAlaAlaAlaAlaHisGlyValArgValLeuGluAsp 160
 DB 761 GGGCCCCGCTCGCGCGCTCGCGAGCTCTCGCGCATGGCGTGGAGGAC 820
 QY 161 GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerLysPheLeuLeuAla 180
 DB 821 GGGGTTAAATTTGCAACAGGGAATTAACCGGTTGCTCCTTTCTATCTTCTGCTGCC 880
 QY 181 LeuLeuSerCysValLeuThrProValSerAlaAlaGluValLysAsnLysSerThrGly 200
 DB 881 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 940
 QY 201 TyrMetValThrAsnAspCysThrAsnAspSerLysThrTrpGlnLeuGlnAlaAlaVal 220
 DB 941 TACATGGTGACCAAGCTGACCAATGATAGTATACCTGGCACTCCAGGCTGCTGCT 1000
 QY 221 LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTrpLys 240
 DB 1001 CTCACGCTCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1060
 QY 241 ProValSerProAsnValAlaValGlnArgProGlyValAlaLeuThrGlnGlyLeuArgThr 260
 DB 1061 CCGGTCTCAGCAATGTGGCGTGCAGAGCCCGGCGGCTTCACGAGGGCTTACGAGCG 1120
 QY 261 HisLeuAspMetValValMetSerAlaThrLeuLeuCysSerAlaLeuTyrValGlyAspLeu 280
 DB 1121 CACATTGACATGGTTGATGTCGCGCACGCTCTGCTCGGCTCTTTACGTGGGGGACCTC 1180
 QY 281 CysGlyGlyValMetLeuAlaAlaGlnMetPheLeuValSerProGlnHisHisTrpPhe 300
 DB 1181 TGGGTGGGGTGATGCTTGCAGCCAGATGTTTATGTTCTGCGCACAGCACCACTGGTTT 1240

QY 301 ValGlnAspCysAsnCysSerIleTyrProGlyThrIleThrGlyHisArgMetAlaTrp 320
 DB 1241 GTGCAAGACTGCAATTTGCTCCATCTACCTGCTGATCCATCACTGGACACCGCATGGCGTGG 1300
 QY 321 AspMetMetMetAsnTrpSerProThrAlaThrMetIleLeuAlaTyrAlaMetArgVal 340
 DB 1301 GACATGATGATGAATCGGTCGCCCGCATCATCATGATCCTGGCGTACCGATGGCGGTC 1360
 QY 341 ProGluValIleIleAspIleIleSerGlyAlaHisTrpGlyValMetPheGlyLeuAla 360
 DB 1361 CCGAGGTGATCATAGACATCATTTGGCGGGCTCATTTGGGGCTCATGTTTGGCTAGCC 1420
 QY 361 TyrPheSerMetGlnGlyAlaTrpAlaLysValValIleLeuLeuLeuAlaAlaGly 380
 DB 1421 TACTTCTCTATGACGAGCGTGGGCAAAAGTCGTTGTCATTTCTTTTGGCGCGCGG 1480
 QY 381 ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr 400
 DB 1481 GTGGAGCGCCCAACCCATACCTTGGGGTTCCTACCGCGCATACCCAGGACCCCTCACC 1540
 QY 401 SerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTrp 420
 DB 1541 GGCATGTTCTCCTTGGTGGCGAGCAAAATCCAGCTCATCAACCAATGGCAGTGG 1600
 QY 421 HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer 440
 DB 1601 CACATCAACCGCACCGCTGAATGCAATGACTCTTTGCACACCGGCTTCTTCGCGTCA 1660
 QY 441 LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg 460
 DB 1661 CTGTTCTACACCCACAGCTTCACTGTCAGGATGTCGCAAGCAATGTCGCGCTCCCGC 1720
 QY 461 SerIleGluAlaPheArgValGlyTrpGlyAlaLeuGlnTyrGluAspAsnValThrAsn 480
 DB 1721 AGTATCGAGCCCTTTCGGTGGATGGCGGCTTACATATATGAGGACCAATGTCCCAAT 1780
 QY 481 ProGluAspMetArgProTyrCysTrpHisTyrProProArgGlnCysGlyValValSer 500
 DB 1781 CCAGAGGATATGAGACCGTATTGTCGCACTACCCACCAAGACAGTGTGTGTAGTCTCC 1840
 QY 501 AlaLysThrValCysGlyProValTyrCysPheThrProSerProValValValGlyThr 520
 DB 1841 GCGAGCTCTGTGTGGCGGCTGCTGTTTCCACCCCGAGCTAGTAGTGGGTAGC 1900
 QY 521 ThrAspArgLeuGlyAlaProThrTyrThrTrpGlyGluAsnGluThrAspValPheLeu 540
 DB 1901 ACCGATAGACTTGGAGCGCCCATTTACACGTTGGGGGAGNATGAGACAGATGCTTCTTA 1960
 QY 541 LeuAsnSerThrArgProProLeuGlySerTrpPheGlyCysThrTrpMetAsnSerSer 560
 DB 1961 TTGAACAGCACTCGACCCAGCGGGTTCATGTTGCGGTGCACTGGATGAATCAATCCACT 2020
 QY 561 GlyTyrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer 580
 DB 2021 GGCTACACCAAGACTTGGCGCGCACCCCTCGCGCATTTAGAGCTAGCTCAATGCGCAGC 2080
 QY 581 ThrAspLeuLeuCysProThrAspCysPheArgLysHisProAspThrThrTyrLeuLys 600
 DB 2081 ATGAGCTTGTGTGCCCCACGAGCTGTTTATGAGAGCATCTCTGATACCACTATACATAA 2140
 QY 601 CysGlySerGlyProTrpLeuThrProArgCysLeuIleAspTyrProTyrArgLeuTrp 620
 DB 2141 TGTGGCTCTGGGCTGGCTCAGCCCAAGGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCT 2200
 QY 621 HisTyrProCysThrValAsnTyrThrIlePheLysIleArgMetTyrValGlyVal 640
 DB 2201 CATTTACCCCTGACAGTAACTATACCATCTTCAAAATAAGGATGATGTGGGGGGGCT 2260
 QY 641 GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp 660
 DB 2261 GAGCAGAGGCTCAGCGCTGCGCAATTTTCACTCGTGGGATGCTGTCACACTTGGAGGAC 2320
 QY 661 ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGluTrpAlaIleLeuPro 680

QY 2561 IleValSerValTrpIleValSerLeuLeuGluAspSerGluThrProIleProThrThrIle 2580
 DB 8021 ATCAAGTCCGTGTGGAGGACCTCTGGAGGACTCAGAAACACCAATTTCCACCAACCAAT 8080
 QY 2581 MetAlaIysAsnGluValPheCysValAspProThrIleValGlyValGlyValAlaAlaArg 2600
 DB 8081 ATGCCAAATATGAGGTGTTCTGGTGGACCCCAACAGGGGGGCAAGAAAGCAGCTCGC 8140
 QY 2601 LeuIleValTrpProAspLeuGlyValArgValCysGluLysMetAlaLeuTrpAspIle 2620
 DB 8141 CTATCGTTTACCTCGACCTCGCGGTGAGGCTCTGGAGAGATGCGCCCTTATGACAT 8200
 QY 2621 ThrGlnLysLeuProGlnAlaValMetGlyAlaSerTrpGlyPheGlnTrpSerProAla 2640
 DB 8201 ACACAAAATCTCTCAGCGCGGTATGGGGCTTCTTATGATTCAGTATTCGCCCGCT 8260
 QY 2641 GlnArgValGluPheLeuLeuLysAlaTrpAlaGluLysAspProMetGlyPheSer 2660
 DB 8261 CAGCGGTAGAGTTCTCTTGAAGATGCGCGGAAAGAGGACCTATGGGTGTTTTCG 8320
 QY 2661 TyrAspThrArgCysPheAspSerThrValThrGluArgAspIleArgThrGluGluSer 2680
 DB 8321 TATGATACCGAGTCTTGTACTCAACGTCACCTGAGAGAGACATCAGGACTGAGGAGTCC 8380
 QY 2681 IleTyrArgAlaCysSerLeuProGluGluAlaHisThrAlaIleHisSerLeuThrGlu 2700
 DB 8381 ATATATGCGGCTGTCTCTTGGCCGAGGAGGCCACACTGCCATACACTCGCTAACTGAG 8440
 QY 2701 ArgLeuTrpValGlyGlyProMetPheAsnSerLysGlyGlnThrCysGlyTrpArgArg 2720
 DB 8441 AGACTTTAGCTGGGAGGGGCTATGTTCAACAGCAGAGGGGCCAACTCGCGGTACAGCGT 8500
 QY 2721 CysArgAlaSerGlyValLeuThrThrSerMetGlyAsnThrIleThrCysTrpValLys 2740
 DB 8501 TGCGCGCCAGCGGGGTGTCCACTAGCATAGTGGGAGACCACTCATCTACGTGAGAA 8560
 QY 2741 AlaLeuAlaAlaCysLysAlaAlaGlyIleAlaProThrMetLeuValCysGlyAsp 2760
 DB 8561 GCCTTAGCGCTGTAAAGCTGAGGATAATCGCGCCACCAATGTGTGTATGCGCGAT 8620
 QY 2761 AspLeuValIleSerGluSerGlnGlyThrGluGluAspGluArgAsnLeuArgAla 2780
 DB 8621 GACTTGGTGTCTATCTCAGAAAGCCAGGGGACCGAGAGGAGCGAGCGAACCTGAGAGCC 8680
 QY 2781 PheThrGluAlaMetThrArgTrpSerAlaProProGlyAspProProArgProGluTrp 2800
 DB 8681 TTCACGAGGCTATGACAGGATATCTGCCCCCTCTGTGACCCCGCCAGACCGAGAT 8740
 QY 2801 AspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaLeuGlyProGlnGly 2820
 DB 8741 GATCTGAGCTGATAACATCTTCTCTCAATGTCTGTGGCGTGGGCCACACAGGC 8800
 QY 2821 ArgArgArgTrpTrpLeuThrArgAspProThrProIleAlaArgAlaTrpGlu 2840
 DB 8801 CGCGCGAGACTACTACCTGACAGAGACCTTACCTCAATGCGCGCGGTGGCGTGGGAA 8860
 QY 2841 ThrValArgHisSerProValAsnSerTrpLeuGlyAsnIleIleGlnTrpAlaProThr 2860
 DB 8861 ACAGTTAGACACTCCCTCTGCAATTCATGGCTGGGAAACATCATCAGTACGCGCCGACC 8920
 QY 2861 IleTrpAlaArgMetValLeuMetThrHisPheSerIleLeuMetAlaGlnAspThr 2880
 DB 8921 ATATGGCTCGCATGCTGATGACACACTTCTCTCTCATCTCATGCTCAAGACAGC 8980
 QY 2881 LeuAspGlnAsnLeuAsnPheGluMetTrpGlyAlaValTrpSerValSerProLeuAsp 2900
 DB 8981 CTGGACAGAACCTCACTTTCAGATGTACGAGCGGTGTACTCCGTGAGTCCCTGGAC 9040
 QY 2901 LeuProAlaIleIleGluArgLeuHisGlyLeuAspAlaPheSerLeuHisThrTrpThr 2920
 DB 9041 CTCCTAGCTATATTAAGGTTACATGGCTTGACGCTTTTCTCTGCACACATACAT 9100
 QY 2921 ProHisGluLeuThrArgValAlaSerAlaLeuArgLysLeuGlyAlaProProLeuArg 2940

DB 9101 CCCACGAACTGACACGGGTGGCTTTCAGCCCTCAGAAAACTTGGGGCGCCACCCCTCAGA 9160
 QY 2941 AlaTrpLysSerArgAlaArgAlaValArgAlaSerLeuIleSerArgGlyGlyArgAla 2960
 DB 9161 GGTGGAGAGCGGGCAGCTGCGTCCAGGGGTCTCTCATCTCCGTCGGCGCGGGGAGCG 9220
 QY 2961 AlaValCysGlyArgTrpLeuPheAsnTrpAlaValLysThrLysLeuLysLeuThrPro 2980
 DB 9221 GCGGTTTGGGTGATATCTCTTCAATTGGCGGTGAACCAAGCTCAAACTCACTCCA 9280
 QY 2981 LeuProGluAlaArgLeuLeuAspLeuSerSerTrpPheThrValGlyAlaGlyGly 3000
 DB 9281 TTGCGGAGAGCGGCTCTCTGATTTATCCAGTGTTCACGCTCGCGCGCGGGGCG 9340
 QY 3001 AspIleTrpHisSerValSerArgAlaArgProArgLeuLeuPheGlyLeuLeu 3020
 DB 9341 GACATTTATCACAGCGGTGCGGTGCGGACCCCGCTTATGCTCTTTGGCCTACTCCTA 9400
 QY 3021 LeuPheValGlyValGlyLeuPheLeuLeuProAlaArg 3033
 DB 9401 CTTTGTAGGGGTAGGCTTTTCTACTCTCCCGCTCGG 9439
 RESULT 4
 AAQ38218
 ID AAQ38218 standard; cDNA to mRNA; 9589 BP.
 XX
 AC AAQ38218;
 XX
 DT 25-MAR-2003 (updated)
 DT 01-JUL-1993 (first entry)
 XX
 DE NANBH virus strain HC-J6 genome cDNA sequence.
 XX
 KW Non A non B hepatitis virus; amplification; HC-J1; HC-J8; plasma; ss.
 XX
 OS Non A, non B hepatitis virus strain HC-J6.
 XX
 FH Key Location/Qualifiers
 CDS 341..942
 FT /*tag= a
 XX
 EP532167-A2.
 XX
 PD 17-MAR-1993.
 XX
 PF 30-JUL-1992; 92EP-0306952.
 XX
 PR 09-AUG-1991; 91JP-0287402.
 PR 05-DEC-1991; 91JP-0360441.
 XX
 PA (IMMO) IMMUNO JAPAN INC.
 XX
 PI Nakamura T, Okamoto H;
 XX
 DR WPI: 1993-087166/11.
 DR N-PSDB; AAR33538.
 XX
 PT Polynucleotide(s), polypeptide(s) and antibodies of NANBH virus -
 PT useful for detecting NANBH, as a vaccine and for screening blood
 PT samples
 XX
 PS Claim 2; Page 27-32; 93pp; English.
 XX
 CC RNA was isolated from the plasma of human patients positive for
 CC NANBH virus (strain HC-J6) and was subjected to reverse transcription
 CC to produce cDNA. The resulting cDNA was amplified by PCR, and
 CC nucleic acid sequences determined by analysis of both clones from the
 CC cDNA library and clones obtd. by PCR amplification (36 clones in total).
 CC The NANBH HC-J6 genome was found to contain an open reading frame
 CC encoding a polypeptide precursor of 3033 amino acid residues.
 CC See also AAQ38172-221.
 CC (Updated on 25-MAR-2003 to correct PN field.)

Db 5801 GCATCCCAAAATTCACACCCCGGGGGCCACTGGCTTCGTTGTCAGTGGCCCTAGTGGGA 5860
 Qy 1841 AlaAlaValGlySerIleGlyLeuGlyLysValLeuValAspIleLeuAlaGlyTyrGly 1860
 Db 5861 GCTGCCGTAGGCAGTATAGCTTAGGTAGGTAGGTAGGTAGGTAGGTAGGTAGGTAGGT 5920
 Qy 1861 AlaGlyIleSerGlyAlaLeuValAlaPheLysIleMetSerGlyGlyLysProSerMet 1880
 Db 5921 CGGGGCATTTCGGGGCTCTCGTCGATTCAGATCATGTCGCGGAGAGCCCTCCATG 5980
 Qy 1881 GluAspValValAsnLeuLeuProGlyIleLeuSerProGlyAlaLeuValValGlyVal 1900
 Db 5981 GAGGATGTCGTAACCTTCGTCGCTGGAATTCCTCTCCGGTCCCTTGGTAGTGGGAGTC 6040
 Qy 1901 IleCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrpMet 1920
 Db 6041 ATCTCGCGGCCCAATTCGCGCCACACGTCGGGACCGGGGAGCGCGCTCCCAATGGATG 6100
 Qy 1921 AsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaProThrHisTyrValThr 1940
 Db 6101 AATAGACTCATTCGCTTGTCTCCAGAGGAATACACGTCGCGCCCAACCTACGTGACG 6160
 Qy 1941 GluSerAspAlaSerGlnArgValThrGlnLeuLeuGlySerLeuThrIleThrSerLeu 1960
 Db 6161 GAGTCGGATCGTCGACGCGTGTGACCAACTACTTGGCTCCCTTACCATTAACACGCTG 6220
 Qy 1961 LeuArgArgLeuHisAsnTrpIleThrGluAspCysProIleProCysGlyGlySerTrp 1980
 Db 6221 CTCAGAAAGACTCACAACTGGATTCAGAGACTGCGCCCACTCCATCGCGCGCTCGTGG 6280
 Qy 1981 LeuArgAspValTrpAspTrpValCysThrIleLeuThrAspPheLysAsnTrpLeuThr 2000
 Db 6281 CTCGCGATGTGGGACTGGTTCGACCATCTTAACAGACTTAAATAATTCGCTGACC 6340
 Qy 2001 SerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTyrLys 2020
 Db 6341 TCCAAATTAATCCAAAGATCGCGGCTCCCTTTGCTCTCTGTCAAAAGGGGTACAAG 6400
 Qy 2021 GlyValTrpAlaGlyThrGlyIleMetThrArgCysProCysGlyAlaAsnIleSer 2040
 Db 6401 GCGGTGGGGCGGACATGGCATCATGACACACGCGTCTCTGCGGCGGCAATATCTCT 6460
 Qy 2041 GlyAsnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTrp 2060
 Db 6461 GGCATGTCGCTGGCTCCATGAGATCACGGGCTTAAGACCTGCATGATATCTG 6520
 Qy 2061 GlnGlyThrPheProIleAsnCysTyrThrGluGlyGlnCysValProLysProAlaPro 2080
 Db 6521 CAGGGGACCTTCTCTATCAATTTGTTACACGAGGGCCAGTCCGTCGCCAAGCCGCGCCA 6580
 Qy 2081 AsnPheLysValAlaIleTrpArgValAlaAlaSerGluTyrAlaGluValThrGlnHis 2100
 Db 6581 AACTTTAAGGTGCGCATCTGAGGGTGGCGGCTCAGAGTACGCGAGGTGACGACAC 6640
 Qy 2101 GlySerTyrHisTyrIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeu 2120
 Db 6641 GGGTCATACCACTACATAACAGGACTCACCACTGATTAACCTGAAGTCCCTGCCACTA 6700
 Qy 2121 ProSerProGluPhePheSerTrpValAspGlyValGlnIleHisArgPheAlaProThr 2140
 Db 6701 CCCTCTCCGAGTTCCTTTCTCGGGTGGACGAGTGCAGATGCCATAGTGTTCGCCCCACA 6760
 Qy 2141 ProLysProPhePheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal 2160
 Db 6761 CCGAAGCGGTTCCTCGGGATGAGGTCTCGTTCGCTGGGTGAGTTCATTTGTCGTC 6820
 Qy 2161 GlySerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThr 2180
 Db 6821 GGGTCCGAGCTCTCTCGGACCTCTGAACCGACACAGACGATTTGATGTCATGCTAACA 6880
 Qy 2181 AspProSerHisIleThrAlaGluThrAlaAlaArgArgLeuAlaArgGlySerProPro 2200
 Db 6881 GATCCATCTCATATCAGCGCGGAGACTGCAGCGGGGTTCAGCGGGGGTTCACCCCA 6940

Qy 2201 SerGluAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThr 2220
 Db 6941 TCCGAGGCAAGCTCTCTCGCGAGCCAGCTATCGGCACCATCGCTCGAGCCACCTGCACC 7000
 Qy 2221 ThrHisGlyLysAlaTyrAspValAspMetValAspAlaAsnLeuPheMetGlyGlyAsp 2240
 Db 7001 ACCCAAGGCAAGCTATGATGTGACATGGTGGATGTAACTGTTCATGGGGGCGAT 7060
 Qy 2241 ValThrArgIleGluSerGlySerLysValValValLeuAspSerLeuAspProMetVal 2260
 Db 7061 GTGACTCGATAGAGTCTGGGTCCAAAGTGGTCTGTTCGGACTCTCTCGACCAATGGTC 7120
 Qy 2261 GluGluArgSerAspLeuGluProSerIleProSerGluTyrMetLeuProLysLysArg 2280
 Db 7121 GAAGAAAGGAGCGACCTTGAGCTTCGATACCATCAGATAATCATGCTCCCAAGAAGAG 7180
 Qy 2281 PheProProAlaLeuProAlaTrpAlaArgProAspTyrAsnProProLeuValGluSer 2300
 Db 7181 TTCCCAACAGCTTTACCGGCTTGGGACGCGCTGATTAACAACCCACCGCTTGTGGAAATCG 7240
 Qy 2301 TrpLysArgProAspTyrGlnProAlaThrValAlaGlyCysAlaLeuProProArg 2320
 Db 7241 TGAAGAAGCCAGATTTACCAACCGGCACTGTTCGGGCTGTGCTCTCCCTCTCTAGG 7300
 Qy 2321 LysThrProThrProProProArgArgArgThrValGlyLeuSerGluAspSerIle 2340
 Db 7301 AAAACCCCGACGCTCCCCCAAGGAGCGCGACAGTGGGCTTAAGTGAAGACTCCATA 7360
 Qy 2341 GlyAspAlaLeuGlnLeuAlaIleLysSerPheGlyGlnProProProSerGlyAsp 2360
 Db 7361 GGAGATGCTCTTCAACAGCTGGCCATTAAGTCTTTGGCCAGCCCCCCCCAAGCGCGAT 7420
 Qy 2361 SerGlyLeuSerThrGlyAlaGlyAlaAspSerGlySerGlnThrProProAspGlu 2380
 Db 7421 TCAGGCTTTTCCAGGGGGCGGGGCTGCGGATTCGCGCAGTCAGACGCTCTCTGATGAG 7480
 Qy 2381 LeuAlaLeuSerGluThrGlySerIleSerSerMetProProLeuGluGlyLeuGly 2400
 Db 7481 TTGGCCCTTTCCGAGACAGGTTCCATCTCTTCATGCCCCCCCTCGAGGGGAGCTTGA 7540
 Qy 2401 AspProAspLeuGluProGluGlnValGluProGlnProProGlnGlyValAla 2420
 Db 7541 GATCCAGACTCGAGCTCGAGCAGTAGAGCCCAACCCCCCCCCCAGGGGGGGTGGCA 7600
 Qy 2421 AlaProGlySerAspSerGlySerTrpSerThrCysSerGluGluAspAspSerValVal 2440
 Db 7601 GCTCCCGGCTCGGACTCGGGTCTCTGTTCTTCTCGAGGAGGAGGACTCCGTCGTG 7660
 Qy 2441 CysCysSerMetSerTyrSerTrpThrGlyAlaLeuIleThrProCysSerProGluGlu 2460
 Db 7661 TGCTGCTCCATGTCATCTCTGACCGGGCTCTAATAACTCTCTTGTAGTCCCGAAGAG 7720
 Qy 2461 GluLysLeuProIleAsnProLeuSerAsnSerLeuLeuArgTyrHisAsnLysValTyr 2480
 Db 7721 GAGAGTTACCGATTAAACCTTTGAGCAACTCCCTGTTCGATATATCAACAAGGTGTAC 7780
 Qy 2481 CysThrThrThrLysSerAlaSerLeuArgAlaLysLysValThrPheAspArgMetGln 2500
 Db 7781 TGATCCACAAAGAGCGCTCCTTAAGGGCTTAAGAGTAACCTTTGTATAGATGCA 7840
 Qy 2501 ValLeuAspSerTyrTyrAspSerValLeuLysAspIleLysLeuAlaSerLysVal 2520
 Db 7841 GTGCTCGACTCTTACTAGACTCAGTCTTAAGGACATTAAGTAGCGGCTCCAAAGTTC 7900
 Qy 2521 ThrAlaArgLeuLeuThrMetGluGluAlaCysGlnLeuThrProProHisSerAlaArg 2540
 Db 7901 ACCGCAAGGCTCTCCATCGAGGAGGCTTGCAGTTAAACCCCAACCTCTTCAAGA 7960
 Qy 2541 SerLysTyrGlyPheGlyAlaLysGluValArgSerLeuSerGlyArgAlaValAsnHis 2560
 Db 7961 TCTAAATATGGGTTTGGGGCTTAAGGAGTCCGAGCTTGTCCGGGAGGGCGCTTAACAC 8020

1101	ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTyrProSerProGlyI	1120
3641	GTCACGCAGATGTACTCCAGTGTCTGAGGGGGACTTAGTAGGGTGGCCAGCCGCTGGG	3700
1121	ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArgAsn	1140
3701	ACTAAATCTTTGGAGCCGTGCACGTGTGGAGCGGTGCACCTGTACTCTGTGTCTACGCGGAAC	3760
1141	AlaAspValIleProAlaArgArgArgGlyAspLysArgGlyAlaValLeuLeuSerProArg	1160
3761	GCTGATGTCTATCCCGCTCGAAGACCGCGGGHCAACACGGGAGCGGTACTCTCCCGAGA	3820
1161	ProLeuSerThrLeuLysGlySerSerGlyGlyProValLeuCysProArgGlyHisAla	1180
3821	CCTCTTTCACCTTGAAGGGGTCTCAGAGGCGCGGTGCTATGCCCCAGGGGCCACGCT	3880
1181	ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle	1200
3881	GTCCGAGTCTTCGGGCACTGTGTGCTCTCGGGCGGTGGCTTAAGTCCATAGATTTCATC	3940
1201	ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro	1220
3941	CCCGTTGAGACACTCGACATCGTCAGCGGTCCCCACCTTTAGTGACAAACGACACCA	4000
1221	ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLys	1240
4001	CCTGTGTGCCCGACACCTATCAGTTCGGGTACTTGCATGCCCGCATGCGCAGTGGAAAG	4060
1241	SerThrLysValProValAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnPro	1260
4061	AGCACCAAGATTCTCTGTCGATATGCTGCTCAGGGGTATAAAGTGTAGTGTCTTAATCCC	4120
1261	SerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerLysAlaHisGlyIleAsnPro	1280
4121	TCAGTGGCTGCCACCTCGGGTTTGGGGCGTACTTGTCTAAGGCATCGCATCAATCCC	4180
1281	AsnIleArgThrGlyValArgThrValThrThrGlyAlaProIleThrTyrSerThrTyr	1300
4181	AACATTAGGACTGGAGTCAGGACTGTGACGACCGGGGCCCATCAGCTACTCCACATAT	4240
1301	GlyLysPheLeuAlaAspGlyCysAlaGlyGlyAlaTyrAspIleIleIleCysAsp	1320
4241	GGCAAAATCTCTCGCCGATGGGGGTGTGGCGGCGGCTACGCATCATCATATGTGAT	4300
1321	GluCysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGlnAla	1340
4301	GAATGCCATGCCGTGGACTCTACACCATCTCTGGCATCGGAACAGTCTCTGATCAAGCA	4360
1341	GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProProGlySerValThr	1360
4361	GAGACAGCTGGGGTCAGACTAATCTGTCTGGCTACAGCTACGCCCTGGGTGAGTGACA	4420
1361	ThrProHisProAsnIleGluValAlaLeuGlyGlnGluGlyGluIleProPheTyr	1380
4421	ACCCCCCAACCAATAGAGAGGTGGCCCTTTGGGCAGAGGGGGAGATCCCTCTCTAT	4480
1381	GlyArgAlaIleProLeuSerTyrIleLysGlyArgHisLeuIlePheCysHisSer	1400
4481	GGGAGGGCGATTCCCTGTCTTACATCAAGGGAGGAAGACATCTGATCTTCTGCCATTCA	4540
1401	LysLysLysCysAspGluLeuAlaAlaLeuArgGlyMetGlyLeuAsnSerValAla	1420
4541	AAGAAAAAGTGTGACGAGCTCGCGCGGCCCTTCGGGGTATGGGTCTGAACCTCAGTGCA	4600
1421	TyrTyrArgGlyLeuAspValSerValIleProThrGlnGlyAspValValValAla	1440
4601	TACTACAGAGGGTTGGAGCTCTCCCGTAATACCAACTCAGGGAGAGCGTAGTGTCTCGCC	4660
1441	ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnVal	1460
4661	ACCGACGCCCTCATGACAGGGTATCTGGGGACTTTGATCTCGTGATCGACTGCAACGTA	4720
1461	AlaValThrGlnValValAspPheSerLeuAspProThrPheThrIleThrGlnIle	1480

4721	DB		GGGGTCACTCAAGTTGTAGACTTCAGTTT	AGACCCCAATTACCATTAAACACACAGATT	4780
1481	QY		ValProGlnAspAlaValSerArgSerGlnArgArgGlyArgThrGlyArgGlyArgLeu	1500	
4781	DB		GTCCCTCAAGACGCTGTCTCAGTAGCAGCGCGGGGTGCGACGGGTAGGGGAAGACTG	4840	
1501	QY		GlyIleTyrArgTyrValSerThrGlyGluArgAlaSerGlyMetPheAspSerValVal	1520	
4841	DB		GGCATTTATAGGTATGTTTCCACTGGTGAGCGAGGCTCAGGATGTGTTGACAGTGTAGTG	4900	
1521	QY		LeuCysGluCysTyrAspAlaGlyAlaAlaTyrTyrGluLeuThrProSerGlnThrThr	1540	
4901	DB		CTCTGTGAGTGTACGACGAGGGGCCGATGGTATGAGCTCACACCATCGAGACCACC	4960	
1541	QY		ValArgLeuArgAlaTyrPheAsnThrProGlyLeuProValCysGlnAspHisLeuGlu	1560	
4961	DB		GTCAGGCTCAGGGCGGTATTTCAACACGCGCTTGCTGTGTGCCAAGACCATCTTGAG	5020	
1561	QY		PheTyrGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThr	1580	
5021	DB		TTTTGGGAGGCAGTTTTCACCGGCTCACACATAGATGCCACTTCTTTCCAAACA	5080	
1581	QY		LysGlnSerGlyGluAsnPheAlaTyrLeuThrAlaTyrGlnAlaThrValCysAlaArg	1600	
5081	DB		AAGCAATCGGGGAAAAATTTCGCATACTTAACAGCCTACCAAGGCTACAGTGTGGCTAGG	5140	
1601	QY		AlaLysAlaProProProSerTyrAspValMetTyrLysCysLeuThrArgLeuLysPro	1620	
5141	DB		GCCAAAGCCCCCCCCCGTCTGCGAGCGTCAATGTGAAGTGTGTGACTCGACTCAAGCCC	5200	
1621	QY		ThrLeuValGlyProThrProLeuLeuTyrArgLeuGlySerValThrAsnGluValThr	1640	
5201	DB		ACACTCGTGGGCCCCACACCTTCTCTGTACCGCTTGGGCTCTGTATTACCAAGAGGTCAAC	5260	
1641	QY		LeuThrHisProValThrLysTyrIleAlaThrCysMetGlnAlaAspLeuGluValMet	1660	
5261	DB		CTCACACATCCGTGACGAAATACATGCCACCTGCATGCAAGCGCACCTTAGGCTCATG	5320	
1661	QY		ThrSerThrTyrValLeuAlaGlyGlyValLeuAlaAlaValAlaAlaTyrCysLeuAla	1680	
5321	DB		ACCAAGCAGCATGGGTCTTGGCAGGGGGAGTCTTGGCGCGGTCTGGCGGCGTATTGGCTGGCG	5380	
1681	QY		ThrGlyCysValCysIleIleGlyArgLeuHisIleAsnGlnArgAlaValAlaAlaPro	1700	
5381	DB		ACCGGGTGTGTTTGCATCATCGGCGCTTGACACATTAAACAGCGAGCGCTGTGTGGCGCG	5440	
1701	QY		AspLysGluValLeuTyrGluAlaPheAspGluMetGluGluCysAlaSerArgAlaAla	1720	
5441	DB		GACAAGGAGGTCTCTATGAGGCTTTTGATGAGATGGAGGAATGTGCTCTAGGCGCGCT	5500	
1721	QY		LeuIleGluGluGlyGlnArgIleAlaGluMetLeuLysSerLysIleGlnGlyLeuLeu	1740	
5501	DB		CTCATTTGAAGAGGGGACGCGATACCGAGATGCTGAAGTCCAAAGTCCAAAGGCTTATTG	5560	
1741	QY		GlnGlnAlaSerLysGlnAlaGlnAspIleGlnProThrValGlnAlaSerTyrProLys	1760	
5561	DB		CAGCAGCTTCCAAACAGCTCAGACATCAACCCACTGTGTGAGGCTTCATGTGCCCAAG	5620	
1761	QY		ValGluGlnPheTyrAlaLysHisMetTyrAsnPheIleSerGlyIleGlnTyrLeuAla	1780	
5621	DB		GTAGAACAATCTGGGCCAAACACATGTGGAATCTCATTTAGCGGCATCCAATACCTCGCA	5680	
1781	QY		GlyLeuSerThrLeuProGlyAsnProAlaValAlaSerMetMetAlaPheSerAlaAla	1800	
5681	DB		GGATATCAACACTGCCCAAGGGAACCTTCAGTAGCTTCCATGATGTGGCGTTCAGTGGCGCC	5740	
1801	QY		LeuThrSerProLeuSerThrSerThrThrIleLeuLeuAsnIleLeuGlyGlyTyrLeu	1820	
5741	DB		CTCACCGAGTCCGCTGTCAACAAAGCACCACTATCTCTTCAACATTTTGGGGGGCTGGCTA	5800	
1821	QY		AlaSerGlnIleAlaProProAlaGlyAlaThrGlyPheValValSerGlyLeuValGly	1840	

1421 TACTTCTATGAGGAGCGTGGCGAAAGTCGTTGTTCATCTCTGTTGGCGCGCGG 1480
 381 ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrGlyArgLeuThr 400
 1481 GTGGACGGCGCACCATACTGTTGGGGGTTTCGCCGCGAGACCAACCGCGCGCTCACC 1540
 401 SerLeuPheAspMetGlyProArgGlnLysAlaGlnLeuValAsnThrAsnGlySerTrp 420
 1541 AGCTTATTTGACATGGGCGCCAGGACGAAATCCAGCTCGTTAAACCAATGGCAGCTG 1600
 421 HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer 440
 1601 CACATCAACGCGACCGCGCTTGAATGCAATGACTCTCTTGACACCGCGCTTATTCGGCTC 1660
 441 LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg 460
 1661 CTGTCTTACACCCACAGCTTCAACTCGTCAGGATGTCGCCAGCGATGTCGCGCTGCCG 1720
 461 SerIleGluAlaPheArgValGlyTrpGlyAlaLeuGlnTyrGluAspAsnValThrAsn 480
 1721 AGTATCGAGGCGCTTCGGGTGGGATGGGCGCGCTTGCAATATGAGGATATGTCACCAAT 1780
 481 ProGluAspMetArgProTyrCysTrpHisTyrProProArgGlnCysGlyValValSer 500
 1781 CCAGAGGATATGAGACCTATGCTGGCAGCTACTCCACCAAGGCGAGTGTGGCGTGTCTCC 1840
 501 AlaIysThrValCysGlyProValTyrCysPheThrProSerProValValValGlyThr 520
 1841 GCGAAGACTGTGTGGCGCCAGTGTACTGTTTCAACCCCGAGCGAGTGTGGCGGACG 1900
 521 ThrAspArgLeuGlyAlaProThrTyrThrTrpGlyGluAsnGluThrAspValPheLeu 540
 1901 ACCGACAGGCTTGGAGCGCCACCTTACACGCTGGGGGAGGATGAGACAGATGCTTCTTA 1960
 541 LeuAsnSerThrArgProProLeuGlySerTrpPheGlyCysThrTrpMetAsnSerSer 560
 1961 TTGAACAGCACTCGACCAACCGCTGGGTCATGGTTTGGCTGCACGTCGATGAATCTTCT 2020
 561 GlyTyrThrIysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer 580
 2021 GGTCTACACCAAGACTTGGCGGCGCACCACTTGGCGTACTAGAGCTGACTTCAACGCGCAG 2080
 581 ThrAspLeuLeuCysProThrAspCysPheArgIysHisProAspThrThrTyrLeuLys 600
 2081 ACGGACCTGTGTGCGCCACCGACTGTTTGGGAAGCATCTGTATACCACTTACCTCAA 2140
 601 CysGlySerGlyProTrpLeuThrProArgCysLeuIleAspTyrProTyrArgLeuTrp 620
 2141 TGGGCTCTGGGCGCTGGCTCACGCCAAGGTGCTGATCGACTACCCCTACAGGCTCTGG 2200
 621 HisTyrProCysThrValAsnTyrThrIlePheLysIleArgMetTyrValGlyVal 640
 2201 CATTACCCCTGACAGTAACTATACCATCTTCAAAATAGAGATGATGTTGGGGGGGTT 2260
 641 GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp 660
 2261 GAGCAGAGGCTCAGCGCTGCATGCAATTTCACTGTGGGATCTGTGCAAAATGGAGGAC 2320
 661 ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGluTrpAlaIleLeuPro 680
 2321 AGACAGAGAGTCAACTGTCTCTCTTGTGCACTCCACCAAGGATGGGCAATTTTACCT 2380
 681 CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisLeuHisGlnAsnIle 700
 2381 TGCTCTTACTCGGACCTGCCCGCTTGTTCGATGCTCTTCCACCTCCACCAAAACATC 2440
 701 ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrLysTyrIleValArgTrp 720
 2441 GTGAGCGTACAAATTCATGATATGGCTATACCTGCGCTCCACAAATATACCTCGGATGG 2500
 721 GluTrpValIleLeuLeuPheLeuLeuAlaAspAlaArgValCysAlaCysLeuTrp 740
 2501 GAGTGGGTAACTCTTATCTCTCTTCTAGCGGAGCGGCTTGGCGCTGTCTTATGG 2560

741 MetLeuIleLeuLeuGlyGlnAlaGluAlaAlaLeuGluLysLeuValIleLeuHisAla 760
 2561 ATGCTCATCTCTTGGCCAGCGCCGAGCAGCAGCTAGAGAAGCTGTCATCTTGCACGCT 2620
 761 AlaSerAlaAlaSerCysAsnGlyPheLeuTyrPheValIlePhePheValAlaAlaTrp 780
 2621 GCGAGCGCAGCTAGCTGCAATGGCTTCTTATATTTTGTCTATCTTTTTCGTCGCTGCTGG 2680
 781 TyrIleLysGlyArgValValProLeuAlaThrTyrSerLeuThrGlyLeuTrpSerPhe 800
 2681 TTACATCAGGGTTCGGGTAGTCCCTTAGCTACTTATTCCTCACTGCGCTGTGCTCTTT 2740
 801 SerLeuLeuLeuAlaLeuProGlnGlnAlaTyrAlaTyrAspAlaSerValHisGly 820
 2741 AGCTACTCTCTCTAGCATTTGCCCAACAGGCTTATGCTTATGACGCACTGTGTGATGGC 2800
 821 GlnIleGlyAlaAlaLeuLeuValMetIleThrLeuPheThrLeuThrProGlyTyrLys 840
 2801 CAGATAGGAGCGGCTCTGCTGGTAAATGATCACTCTCTTTTACTCTTCAACCCCGGTTAAG 2860
 841 ThrLeuLeuSerArgPheLeuTrpTyrLeuCysTyrLeuLeuThrLeuGlyGluAlaMet 860
 2861 ACCCTTCTCAGCGGTTTTTGTGGTGTGCTATCTTCTGACCTTGGGGAGGATATG 2920
 861 ValGlnGluTrpAlaProProMetGlnValArgGlyGlyArgAspGlyIleIleTrpAla 880
 2921 GTCCAGAGTGGGACCACTATGACAGGTGGCGGTGGCGTGTGATGGCATATATGGGCC 2980
 881 ValAlaIlePheTyrProGlyValValPheAspIleThrIysTrpLeuLeuAlaValLeu 900
 2981 GTCCCATATTTACCCAGGTGTGTGTTTGAATTAACCAAGTGGCTCTTGGCGGTGCTT 3040
 901 GlyProAlaTyrLeuLeuLysGlyAlaLeuThrArgValProTyrPheValArgAlaHis 920
 3041 GGGCTTGTCTTACCTCTCTAAAGGTGCTTGGCGCGGTGCGACTTCTGTCAGGGCTCAC 3100
 921 AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyGlyArgTyrValGlnMet 940
 3101 GCTCTACTGAGGATGTCACCATGCAAGGCACTCTCGCGGGGCGAGTACGTCCAGATG 3160
 941 AlaLeuLeuAlaLeuGlyArgTrpThrGlyTyrIleTyrAspHisLeuThrProMet 960
 3161 GCGCTACTAGCCCTTGGCAGGTGGACTGGCACTTATATCTATGACCACTTCAACCTTATG 3220
 961 SerAspTrpAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIleIlePhe 980
 3221 TCGGATTTGGGCTGCTAGTGGCTTGGCGGACCTGCGGCTGCGCTTGGAGCTTATCATCTTC 3280
 981 SerProMetGluLysLysValIleValTrpGlyValAlaGluThrAlaLysGlyAspIle 1000
 3281 AGTCCGATGGAGAAAGTCAATTTGTCGGGAGCGGAGACAGCTGCTTGTGGGACATTT 3340
 1001 LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp 1020
 3341 TTACACGCACTTCCCGTGTCCGCGCACTTGGTGGGAGGTCTCTTGTGGCCAGCTGAT 3400
 1021 GlyTyrThrSerLysGlyTrpSerLeuLeuAlaProIleThrAlaTyrAlaGlnGlnThr 1040
 3401 GGTATATACCTTCCAAAGGGGTGAGTCTTCTCGCCCCCATCACTGCTTACGCCCGACAGCA 3460
 1041 ArgGlyLeuLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGluGlnAla 1060
 3461 CGTGGCTTTTGGCACCATAGTGGTGGAGCATGCGGGGCGGACAGACAGACAGCT 3520
 1061 GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly 1080
 3521 GGGGAAATTCAGGTCTCTGTCACAGTCACTCAGTCTCTTCTTCTCGGAAATCATCTTCGGG 3580
 1081 ValLeuTrpThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro 1100
 3581 GTTTTGTGGACTGTCTACCATGGAGCTGGCAAGACTCTTGGCGCGCTCTCAGCGGGTCCG 3640

Qy 2941 AlaTrpLysSerArgAlaAlaValArgAlaSerLeuLeuSerArgGlyGlyArgAla 2960
 Db 9161 GCGTGAAGAGCGCGGACGTCAGTCAGGCGTCCCTCATCTCCCGTGGGGGAGAGCG 9220
 Qy 2961 AlaValCysGlyArgTyrLeuPheAsnTrpAlaValLysThrLysLeuLysLeuThrPro 2980
 Db 9221 GCGCTTTGGCGTGCATATCTCTTCAATGGGCGGTGAAGACCAAGCTCAAACTCACTCCA 9280
 Qy 2981 LeuProGluAlaArgLeuLeuAspLeuSerSerTyrPheThrValGlyAlaGlyGlyGly 3000
 Db 9281 TTGCGGAGAGCGGCTCTCTGGATTATCCAGCTGGTTCACCGTCGGCGCGCGGGGGC 9340
 Qy 3001 AspIleTyrHisSerValSerArgAlaArgProArgLeuLeuLeuPheGlyLeuLeuLeu 3020
 Db 9341 GACATTATACAGCGTGTGCGGTGCGCGACCGCGCTTATGTCTTTGGCTTACTCTTA 9400
 Qy 3021 LeuPheValGlyValGlyLeuPheLeuLeuProAlaArg 3033
 Db 9401 CTTTGTAGGGTAGGCCTTTTCTACTCCCGCTCGG 9439

RESULT 3

AAF23486
 ID AAF23486 standard; DNA; 9711 BP.

AAF23486;

21-MAR-2001 (first entry)

Infectious Hepatitis C virus 2a genotype.

GBV-B; hepatitis C virus; HCV; vaccine; ds.

Hepatitis C virus.

WO200075337-A1.

14-DEC-2000.

02-JUN-2000; 2000WO-US15293.

04-JUN-1999; 99US-0137694.

(USSH) US DEPT HEALTH & HUMAN SERVICES.

Bukh J, Yanagi M, Emerson SU, Purcell RH;

WPI; 2001-091214/10.

New infectious nucleic acids of the GB virus-B clone, useful for indirectly studying the molecular properties of hepatitis C virus (HCV) and in developing vaccines and therapeutics for HCV.

Disclosure; Page 78-82; 96pp; English.

The present invention relates to GB virus-B. The nucleic acid molecules of the invention are useful for indirectly studying the molecular properties of hepatitis C virus (HCV). The infectious nucleic acid sequence of the GB virus-B clone and the HCV/GBV-B chimeras may be used in the development of vaccines and therapeutics for HCV.

Sequence 9711 BP; 1989 A; 2851 C; 2668 G; 2203 T; 0 other;

Alignment Scores:

Pred. No.:	0	Length:	9711
Score:	16106.00	Matches:	3031
Percent Similarity:	99.93%	Conservative:	0
Best Local Similarity:	99.93%	Mismatches:	2
Query Match:	99.91%	Indels:	0
DB:	22	Gaps:	0

US-09-980-559-2 (1-3033) x AAF23486 (1-9711)

Qy 1 MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgProGln 20

Db	341	ATGAGCACAAATCTTAACCTCAAGAAAAACCAAGAAACCAACCAACCTGCGCCCAA	400
Qy	21	AspValLysPheProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArg	40
Db	401	GAGCTTAAGTTTCCGGCGCGCGCCAGATCGTTGGCGGAGTATACTTGTTGGCGCGCAGG	460
Qy	41	GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly	60
Db	461	GGCCCCAGGTTGGGTGGCGGCGACAGAGAACTTCGGAGCGGTCCAGCCACGTTGGA	520
Qy	61	ArgArgGlnProLysAspArgArgSerThrGlyLysSerTrpGlyLysProGly	80
Db	521	AGCGCGCAGCCATCCCTAAAGATCGCGCTCCACTGGCAATCTCTGGGGAACCAAGGA	580
Qy	81	TyrProTrpProLeuTyrGlyAsnGluGlyLeuGlyTyrAlaGlyTrpLeuLeuSerPro	100
Db	581	TACCCCTGGCCCCATACCGGGAATGAGGGAATCGGCTGGGACAGATGGCTCTCTGTC	640
Qy	101	ArgGlySerArgProSerTrpGlyProAsnAspProArgHisArgSerArgAsnValGly	120
Db	641	CGAGGTTCCCGTCCCTCTTGGGGCCCCCAATGACCCCGCGCATAGGTTCGCAACGTGGT	700
Qy	121	LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProValVal	140
Db	701	AAGGTCATCGATACCTAACGTGCGGCTTTGGCGACCTCATGGGTACATCCCTGCTGTG	760
Qy	141	GlyAlaProLeuGlyGlyValAlaArgAlaLeuAlaHisGlyValArgValLeuGluAsp	160
Db	761	GGCGCCCGCTCGCGCGCGCTCGCAGAGCTCTCGCGCATGGCGTGAGAGTCTCTGGAGG	820
Qy	161	GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla	180
Db	821	GGGGTTAATTTGCAACAGGGAATTTACCCGGTGTCTCTTTCTATCTCTTCTGTGGCC	880
Qy	181	LeuLeuSerCysIleThrThrProValSerAlaAlaGluValLysAsnIleSerThrGly	200
Db	881	CTGCTCTCTGCATCACACCCCGGTCTCCGCTGCCGAAGTGAAGAACAATCAGTACCGGC	940
Qy	201	TyrMetValThrAsnAspCysThrAsnAspSerIleThrTrpGlnLeuGlnAlaVal	220
Db	941	TACATGGTGACTAACGACTGCACCAATGACAGCATTTACCTGGCAGCTCCAGGCTGTG	1000
Qy	221	LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTrpIle	240
Db	1001	CTCCACGTCCTCCGGGTGCGTCCGTCGAGAAAGTGGGGAATGCATCTCAGTCTGGATA	1060
Qy	241	ProValSerProAsnValAlaValGlnArgProGlyAlaLeuThrGlnGlyLeuArgThr	260
Db	1061	CCGGTCTCACCGAATGTGGCGGTGACGCGCGCGCGCTCACGACGAGGCTTGGCGAGC	1120
Qy	261	HisIleAspMetValValMetSerAlaThrLeuCysSerAlaLeuTyrValGlyAspLeu	280
Db	1121	CACATCGCATGGTTGTGATGTCGCGCAGCTCTGCTCTGCTCTACGCTGGGGGACCTC	1180
Qy	281	CysGlyGlyValMetLeuAlaGlnMetPheIleValSerProGlnHisTrpPhe	300
Db	1181	TGCGGGGGGTGATGTCGAGCCCAATGTTCAITGTCTCGCCGAGACCACTGGGTT	1240
Qy	301	ValGlnAspCysAsnCysSerIleTyrProGlyThrIleThrGlyHisArgMetAlaTrp	320
Db	1241	GTCCAAGACTGCNAATTTGCTCCATCTACCTGGTACCATCTACGTGACACCGCATGGC	1300
Qy	321	AspMetMetMetAsnTrpSerProThrAlaThrMetIleLeuAlaTyrAlaMetArgVal	340
Db	1301	GACATGATGAATGATGTCGCGCGCGCTACCGGTACCATGATCTTGGCGTACGCGATG	1360
Qy	341	ProGluValIleLeuAspIleLeuSerGlyAlaHisTrpGlyValMetPheGlyLeuAla	360
Db	1361	CCCGAGGTCAATATAGACATCATTAGCGGGCTCATTTGGGGCGGTCAATTTGGCTGGCC	1420
Qy	361	TyrPheSerMetGlnGlyAlaTrpAlaLysValValIleLeuLeuLeuAlaGly	380

6941 TCCGAGCAAGCTCTCCGCGAGCCAGCTATCGCCACCATCGCTGCGAGCCACCTGCACC 7000
 2221 ThrHisGlyLysAlaTyrAspValAspMetValAspAlaAsnLeuPheMetGlyGlyAsp 2240
 7001 ACCCAGCGCAAGCCTATGATGTGGACATGGTGGATGCTAACCTGTTATGGGGGGCGAT 7060
 2241 ValThrArgIleGluSerGlySerLysValValValLeuAspSerLeuAspProMetVal 2260
 7061 GTGACTCGGATAGATCTGGGTCAAAGTGGTCTGTTCTGACTCTCTCGACCCCAATGGTC 7120
 2261 GluGluArgSerAspLeuGluProSerIleProSerGluTyrMetLeuProLysLysArg 2280
 7121 GAAGAAAGGAGCGACCTTGAGCTTCGATACCATCAGATACATGCTCCGCCCAAGAGAGG 7180
 2281 PheProAlaLeuProAlaTyrAlaArgProAspTyrAsnProProLeuValGluSer 2300
 7181 TTCCACACAGCTTTACCGGCTGGGCGAGCGCTGATTAACACCCACCGCTTGTTGGAAATCG 7240
 2301 TrpLysArgProAspTyrGlnProAlaThrValAlaGlyCysAlaLeuProProArg 2320
 7241 TGGNAAAGGCCAGATTACCAACCGGCCACTGTTGGGGCTGTGCTCTCCCTCCCTTAGG 7300
 2321 LysThrProThrProProArgArgArgThrValGlyLeuSerGluAspSerIle 2340
 7301 AAAACCCGAGCTCCGCCCAAGAGGCGCGGACAGTGGGCTTAAGTGAGGACTCCATA 7360
 2341 GlyAspAlaLeuGlnGlnLeuAlaIleLysSerPheGlyGlnProProProSerGlyAsp 2360
 7361 GGAGATGCCCTTCAACAGCTGGCCATTAAAGTCTTTGGGCCAGGCCCCCAAGCGCGAT 7420
 2361 SerGlyLeuSerThrGlyAlaGlyAlaAlaAspSerGlySerGlnThrProProAspGlu 2380
 7421 TCAGGCCCTTCCAGCGGGCGGGCGCTGCCGATTCGGCAGTCAGACGCTCCTGATGAG 7480
 2381 LeuAlaLeuSerGluThrGlySerIleSerMetProProLeuGluGlyGluLeuGly 2400
 7481 TTGGCCCTTTCCGAGACAGGTTCCATCTCTCCATGCCCCCCCCCTCGAGGGGGAGCTTGG 7540
 2401 AspProAspLeuGluProGluGlnValGluProGlnProProProGlnGlyGlyValAla 2420
 7541 GATCCAGACCTGGAGCCTGAGCAGTAGAGCCCAACCCCCCCCCAGGGGGGGGGTGGCA 7600
 2421 AlaProGlySerAspSerGlySerThrCysSerGluGluAspAspSerValVal 2440
 7601 GCTCCCGGCTCGGACTCGGGTCTGCTGCTACTTCTCCGAGGAGGAGGACTCCGCTGCTG 7660
 2441 CysCysSerMetSerTyrSerTyrThrGlyAlaLeuLeuThrProCysSerProGluGlu 2460
 7661 TGCTGCTCCATGTCTATCTCTGACCGGGCTCTAATACTCTGTTAGTCCCGAAGAG 7720
 2461 GluLysLeuProIleAsnProLeuSerAsnSerLeuLeuArgTyrHisAsnLysValTyr 2480
 7721 GAGAAGTTACCGAVTAACCCCTTGAGCAACTCCCTGTTGGCATATCAACAAGGTGTAC 7780
 2481 CysThrThrThrLysSerAlaSerLeuArgAlaLysLysValThrPheAspArgMetGln 2500
 7781 TGTACCACAAAGAGCGCTCTACTAAGGGCTAAAGAGGTAACTTTTGATAGGATGCA 7840
 2501 ValLeuAspSerTyrTyrAspSerValLeuLysAspIleLysLeuAlaAlaSerLysVal 2520
 7841 GTGCTCGACTCCTACTACGACTCAGTCTTAAAGGACATTAAGCTAGCGGCTCCCAAGTC 7900
 2521 ThrAlaArgLeuLeuThrMetGluAlaCysGlnLeuThrProProHisSerAlaArg 2540
 7901 ACCGCAAGGCTCCTCACCATCGAGGAGGCTTGCACAGTTTAAACCCCAACCCCTTTCGAAGA 7960
 2541 SerLysTyrGlyPheGlyAlaLysGluValArgSerLeuSerGlyArgAlaValAsnHis 2560
 7961 TCTAATATGGGTTGGGGCTTAAGAGGTCCGAGCTTGTCCGGGAGGGCGCTTAACCCAC 8020
 2561 IleLysSerValTrpLysAspLeuLeuGluAspSerGluThrProIleProThrThrIle 2580
 8021 ATCAAGTCCGTGGAGGAGCTCTCTGGAGGACTCAGAAACACCAATTTCCCAACCAACATT 8080

2581 MetAlaLysAsnGluValPheCysValAspProThrLysGlyGlyLysLysAlaAlaArg 2600
 8081 ATGGCCAAAATAGAGTGTTCGCTGGACCCCAAGGGGGGCAAGAAACAGCTGC 8140
 2601 LeuIleValTyrProAspLeuGlyValArgValCysGluLysMetAlaLeuTyrAspIle 2620
 8141 CTTATCGTTTACCTGACCTCGCGCTCAGGGTCTCGAGAGATGGCCCTTTATGACATT 8200
 2621 ThrGlnLysLeuProGlnAlaValMetGlyAlaSerTyrGlyPheGlnTyrSerProAla 2640
 8201 ACACAAAATCTTCCTCAGCGGTGATGGGGCTTCTTATGGATTCAGTATTTCCCGCT 8260
 2641 GlnArgValGluPheLeuLeuLysAlaTrpAlaGluLysLysAspProMetGlyPheSer 2660
 8261 CAGCGGTAGAGTTCTCTTGAAGCATGGCGGGAAGAGAGCCCTATGGGTTTTTCG 8320
 2661 TyrAspThrArgCysPheAspSerThrValThrGluArgAspIleArgThrGluLys 2680
 8321 TATGATACCCGATGCTTTCGACTCAACCGCTCACTGAGAGAGATCAGGACTGAGGAGTCC 8380
 2681 IleTyrArgAlaCysSerLeuProGluGluAlaHisThrAlaIleHisSerLeuThrGlu 2700
 8381 ATATATCGGGCTTGTCTTCCCGAGAGGCGCCACACTGCCATACACTCGCTTAACCTGAG 8440
 2701 ArgLeuTyrValGlyGlyProMetPheAsnSerLysGlyGlnThrCysGlyTyrArgArg 2720
 8441 AGACTTTACGTTGGAGGGCTTATGTTCAACAGCAAGGGCCAAACCTCGCGGTACAGGCT 8500
 2721 CysArgAlaSerGlyValLeuThrThrSerMetGlyAsnThrIleThrCysTyrValLys 2740
 8501 TGCCGCGCCAGGGGGTGTCTCACCCTAGCATGGGGAACACCATCATCATGCTACGTGAAA 8560
 2741 AlaLeuAlaAlaCysLysValAlaGlyIleAlaProThrMetLeuValCysGlyAsp 2760
 8561 GCCTTAGCGCTGTAAAGCTGAGGAGTAATCGCCGCCACATGCTGTTATGCGCGAT 8620
 2761 AspLeuValValIleSerGluSerGlnGlyThrGluGluAspGluArgAsnLeuArgAla 2780
 8621 GACTTGTGTTGTCATCTCAGAAAGCCAGGGGAGCGAGGAGGAGCGGAACTGAGAGCC 8680
 2781 PheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProArgProGluTyr 2800
 8681 TCCAGCGAGCTATGACAGGATTTCTGCGCTCTCTGCTGACCCCGCCAGCGGAGTAT 8740
 2801 AspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaLeuGlyProGlnGly 2820
 8741 GATCTGGAGCTGATAACATCTTGTCTCTCAATGTGTCTGTGGCGCTGGGCCCAAGGC 8800
 2821 ArgArgArgTyrTyrLeuThrArgAspProThrThrProIleAlaArgAlaAlaTrpGlu 2840
 8801 CGCCGAGATACCTACCTGACGAGACCTTACCCTCCAACTCCGCTGGGCTGCTGGAA 8860
 2841 ThrValArgHisSerProValAsnSerTrpLeuGlyAsnIleIleGlnTyrAlaProThr 2860
 8861 ACAGTTAGACATCCCTCTGTAATTCATGCTGGGAAACATCATCCAGTACGCCCGAC 8920
 2861 IleTrpAlaArgMetValLeuMetThrHisPhePheSerIleLeuMetAlaGlnAspThr 2880
 8921 ATATGGGCTCGCATGGTCTGATGACACATCTTCTTCCATTTCTATGCTCAGACACG 8980
 2881 LeuAspGlnAsnLeuAsnPheGluMetTyrGlyValaValTyrSerValSerProLeuAsp 2900
 8981 CTGGACAGCAACCTCAACTTTGAGATGATACGAGCGGTGTACTCCGAGTCTCTGGAC 9040
 2901 LeuProAlaIleIleGluArgLeuHisGlyLysAspAlaPheSerLeuHisThrThr 2920
 9041 CTCCAGCTATAATGAAAGGTTACATGGGCTTTCAGCGCTTTTCTCTGCAACATACACT 9100
 2921 ProHisGluLeuThrArgValAlaSerAlaLeuArgLysLeuGlyAlaProProLeuArg 2940
 9101 CCCACGAACTGACCGGGTGGCTTACGCTTACGAAACCTTGGGGCGCCACCCCTCAGA 9160

QY 1481 ValProGlnAspAlaValSerArgSerGlnArgArgGlyArgThrGlyArgGlyArgLeu 1500
DB 4781 GTCCCTCAAGACGCTGTCTCAGCTAGCCAGCGCGGGTCCGACGGGTAGGGGAAGACTG 4840
QY 1501 GlyIleTyrArgTyrValSerThrGlyGluArgAlaSerGlyMetPheAspSerValVal 1520
DB 4841 GGCATTATAGGTATGTTTCCATGGTGAGCGAGCCCTCAGGAATGTTTGACAGTGTAGTG 4900
QY 1521 LeuCysGluCysTyrAspAlaGlyAlaAlaTyrTyrGluLeuThrProSerGlnThrThr 1540
DB 4901 CTCTGTGAGTGTACGACGACGCGGGCGCATGTATGAGCTCACCATCGAGACACC 4960
QY 1541 ValArgLeuArgAlaTyrPheAnThrProGlyLeuProValCysGlnAspHisLeuGlu 1560
DB 4961 GTCAGGCTCAGGCGGTATTTCAACACGCGCGTTCCTGTGTGCCAAGACCATCTTGAG 5020
QY 1561 PheTyrGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThr 1580
DB 5021 TTTTGGGAGCGAGTTTTCACCGCGCTCACACATAGATGCCCACTTCTCTTCCCAACA 5080
QY 1581 LysGlnSerGlyGluAnPheAlaTyrLeuThrAlaTyrGlnAlaThrValCysAlaArg 1600
DB 5081 AAGCAATCGGGGGGAAATTTCCGATACCTTAACAGCTTACCGGCTCAGTGTGCGTAGG 5140
QY 1601 AlaLysAlaProProSerTrpAspValMetTrpLysCysLeuThrArgLeuLysPro 1620
DB 5141 GCCAAGCCCGCCCGCTCTGGGAGCTCATGTGGAAGTGTGTGACTCAAGCCC 5200
QY 1621 ThrLeuValGlyProThrProLeuLeuTyrArgLeuGlySerValThrAnGluValThr 1640
DB 5201 ACATCGTGGGCGCCACACCTCTCTGTATCCGCTTGGGCTCTGTATCCACAGAGGTCA 5260
QY 1641 LeuThrHisProValThrLysTyrIleAlaThrCysMetGlnAlaAspLeuGluValMet 1660
DB 5261 CTCACATCCCGTGAGCAATACATCGCACCTGTATGCAAGCCGACCTTGAGGTCA 5320
QY 1661 ThrSerThrTrpValLeuAlaGlyGlyValLeuAlaAlaValAlaAlaTyrCysLeuAla 1680
DB 5321 ACCAGCACATGGGTCTTGGCAGGGGAGTCTTGGCGCGCTCGCGCGTATTGCTGGCG 5380
QY 1681 ThrGlyCysValCysIleIleGlyArgLeuHisIleAsnGlnArgAlaValAlaPro 1700
DB 5381 ACCGGGTGTGTTTGCATCATCGCGCGCTTGCATTAACACAGCGAGCGCTGTGCGCG 5440
QY 1701 AspLysGluValLeuTyrGluAlaPheAspGluMetGluGluCysAlaSerArgAlaAla 1720
DB 5441 GACAGAGGTCTCTATAGGCTTTTGTATGAGATGGAGAAATGTGCTCTAGGGCGGT 5500
QY 1721 LeuIleGluGlyGlnArgIleAlaGluMetLeuLysSerLysIleGlnGlyLeuLeu 1740
DB 5501 CTCAATTGAGAGGGGCGAGTAGCCGAGATGCTGAGTCCAAAGATCCAAAGCTTATTG 5560
QY 1741 GlnGlnAlaSerLysGlnAlaGlnAspIleGlnProThrValGlnAlaSerTrpProLys 1760
DB 5561 CAGCAAGCTTCCAAACAAGCTCAAGCATACAAACCCACTGTGCAGGCTTCATGGCCCAAG 5620
QY 1761 ValGluGlnPheThrAlaLysHisMetTrpAsnPheIleSerClyIleGlnTyrLeuAla 1780
DB 5621 GTAGAACAATTCGGGCGCAACACATGTGGAATCTTATAGCGGCATCAATACCTCGCA 5680
QY 1781 GlyLeuSerThrLeuProGlyAnProAlaValAlaSerMetMetAlaPheSerAlaAla 1800
DB 5681 GGACTATCAACACTGCCAGGAACCTGCGAGTAGCTTCCATGATGCGGTTCAGTGGCGCC 5740
QY 1801 LeuThrSerProLeuSerThrSerThrThrIleLeuLeuAsnIleLeuGlyTyrLeu 1820
DB 5741 CTCACAGTCCGCTGTCAACAGCACCACTATCTCTCAACATTTTGGGGGCTGGCTA 5800
QY 1821 AlaSerGlnIleAlaProProAlaGlyAlaThrGlyPheValValSerGlyLeuValGly 1840
DB 5801 GCATCCCAATTCACACCCCGGGGGCACCTGGGCTTGTGTAGTGGCTTAGTGGGA 5860
QY 1841 AlaAlaValGlySerIleGlyLeuGlyLysValLeuValAspIleLeuAlaGlyTyrGly 1860

DB 5861 GCTCGCGTAGGAGCATATAGCTTAGGTAGGTAGTGGACATCTCGCAGGATGTGT 5920
QY 1861 AlaGlyIleSerGlyAlaLeuValAlaPheLysIleMetSerGlyGluLysProSerMet 1880
DB 5921 GCGGGCATTTCCGGGCTCTCGTCGCAITCAAGATCATGTCTGGCGAGAGCCCTCATG 5980
QY 1881 GluAspValValAsnLeuLeuProGlyIleLeuSerProGlyAlaLeuValValGlyVal 1900
DB 5981 GAGGATGCTCAACTGCTGCTGGAATCTGTCTCCGGGTGCTTGGTAGTGGAGTC 6040
QY 1901 IleCysAlaAlaIleLeuArgHisValGlyProGlyGluGlyAlaValGlnTrpMet 1920
DB 6041 ATCTGCGCGCCATTTCTGCGCGCACAGCTGGGACCGGGGGAAGCGCGCTCCAATGGATG 6100
QY 1921 AsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaProThrHisTyrValThr 1940
DB 6101 AATAGACTCATTTGCTTTCCAGAGGAATCACTGCGCCGCCACCCACTACGTGACG 6160
QY 1941 GluSerAspAlaSerGlnArgValThrGlnLeuLeuGlySerLeuThrIleThrSerLeu 1960
DB 6161 GAGTGGATGCTGCGAGGCTGTGACCCAACTACTTGGCTCTCTTACCATTAACAGCTG 6220
QY 1961 LeuArgArgLeuHisAsnTrpIleThrGluAspCysProIleProCysGlyGlySerTrp 1980
DB 6221 CTCAGAGACTCCACAACTGGATTACTGAGGACTGCCCATCCCATCGCGGCTCGTG 6280
QY 1981 LeuArgAspValTrpAspTrpValCysThrIleLeuThrAspPheLysAsnTrpLeuThr 2000
DB 6281 CTCGCCGATGTGTGGAGCTGGGTTTGCACCACTCTTAACAGACTTTTAAATAATTTGG 6340
QY 2001 SerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTyrLys 2020
DB 6341 TCCAAATTTATCCAAAGATGCGCGGCTCCCTCTCTGTCTCTGTCAAAAGGGGTACAAG 6400
QY 2021 GlyValTrpAlaGlyThrGlyIleMetThrThrArgCysProCysGlyAlaAsnIleSer 2040
DB 6401 GGCATGTGGCGCGCACTGGCATCATGACCAACAGGTGTCTTGGCGGCCAATATCTCT 6460
QY 2041 GlyAnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTrp 2060
DB 6461 GGCATATGCGCTTGGCTTCCATGAGAATCACGGGCGCTTAAGACCTGCAATGATCTG 6520
QY 2061 GlnGlyThrPheProIleAsnCysTyrThrGluGlyGlnCysValProLysProAlaPro 2080
DB 6521 CAGGGGACCTTCTCTATCAATTTTACACGAGGGCCAGTGTGCGGAAACCCCGCCCA 6580
QY 2081 AsnPheLysValAlaIleTrpArgValAlaAlaSerGluTyrAlaGluValThrGlnHis 2100
DB 6581 AACTTTAAGTCCCATCTGGAGGGTGGCGGCTCAGAGTACGCGAGGTGACGAGCAC 6640
QY 2101 GlySerTyrHisTyrIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeu 2120
DB 6641 GGGTCATACCACTACATAACAGGACTCACCACTGATTAACCTTGAAGTCTCCCTGCCAAC 6700
QY 2121 ProSerProGluPhePheSerTrpValAspGlyValGlnIleHisArgPheAlaProThr 2140
DB 6701 CCTCTCCGAGTCTTCTTCTGGGTGGACGAGTGCAGATCCATAGGTTCGCCCCACA 6760
QY 2141 ProLysProPhePheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal 2160
DB 6761 CCGAAGCGTTTTTCCGGGATGAGTCTGCTTCTGCTGGTGGCTTAAATTTCAITTTGTCG 6820
QY 2161 GlySerGlnLeuProCysAspProGluProLysThrAspValLeuMetSerMetLeuThr 2180
DB 6821 GGGTCCCAAGCTTCTTTCGACCTCGAACCCGACAGACGATATGATGTGCTCAATGCTAAC 6880
QY 2181 AspProSerHisIleThrAlaGluThrAlaAlaArgArgLeuAlaArgGlySerProPro 2200
DB 6881 GATCCATCTCATATCACGCGGAGACTGACGCGCGGCTTTAGCGGGGGGTACCCCCCA 6940
QY 2201 SerGluAlaSerSerAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThr 2220

2561	ATGCTCATCTTGTGGCCAGGCCGAGCAGACACTAGAGAACTGGTCACTCTTGACACGCT	2620
761	AlaSerAlaAlaSerCysAsnGlyPheLeuTyrrPheValIlePhePheValAlaIaIaTrp	780
2621	GCAGCGCAGCTAGCTGCAATGGCTTCCTATATTTTGTTCATCTTTTTCGTGCTGCTGG	2680
781	TyrIleLysGlyArgValValProLeuAlaIaThrTyrSerLeuThrGlyLeuTrpSerPhe	800
2681	TACATCAAGGGTCGGGTAGTCCTTAGCTACTATTCCTCCTCAGTCGGCTGGTGCCTTT	2740
801	SerLeuLeuLeuAlaLeuProGlnGlnAlaTyrrAlaTyrrAspAlaSerValHisGly	820
2741	AGCTTACTCTCTAGCATTCGCCCAACAGGCTTATGCTTATGACGCATCTGTGCATGCC	2800
821	GlnIleGlyAlaAlaLeuLeuValMetIleThrLeuPheThrLeuThrProGlyTyrLys	840
2801	CAGATAGGACGGCTCTGCTGGTAATGATCACTCTCTTTTACTCTCACCCCGGGTATAAG	2860
841	ThrLeuLeuSerArgPheLeuTrpTrpLeuCysTyrLeuLeuThrLeuGlyGluAlaMet	860
2861	ACCTTCTCAGCGGGTTTGTGTGTGTGTGCTATCTTCTGACCCCTGGGGGAAGCTATG	2920
861	ValGlnGluTrpAlaProProMetGlnValArgGlyGlyArgAspGlyIleIleTrpAla	880
2921	GTCCAGAGTGGGCACCACTATGCAAGGTGCGCGGTGGCGTATGGCATCATATGGGGC	2980
881	ValAlaIlePheTyrProGlyValValPheAspIleThrLysTrpLeuLeuAlaValLeu	900
2981	GTCCGCATATTTACCCAGGTGTGGTGTGTGTGACATPAACCAAGTGGCTCTTGGCGGTCTT	3040
901	GlyProAlaTyrrLeuLeuLysGlyAlaLeuThrArgValProTyrPheValIaArgAlaHis	920
3041	GGGCGCTGCTTACCTCTCAAAAGGTGCTTTGACGCGGTGCGGTACTCTGTCGAGGGCTCAC	3100
921	AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyGlyArgTyrValGlnMet	940
3101	GCTTACTAGGATGTGCACCATGGCAAGGCATCTCGCGGGGGGAGTAGCTGCCAGATG	3160
941	AlaLeuLeuAlaLeuGlyArgTrpThrGlyThrTyrIleTyrAspHisLeuThrProMet	960
3161	GGGCTACTAGCCCTTGGCAGGTGGACTGGCACTTACATCTATGACCACCTCACCCCTATG	3220
961	SerAspTrpAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIleIlePhe	980
3221	TCGGATTGGCTGCTAGTGGCTTGGCGGACCTGGCGGTGCGCGTTGAGCTTATCATCTTC	3280
981	SerProMetGluLysLysValIleValTrpGlyAlaGluThrAlaIaCysGlyAspIle	1000
3281	AGTCCCATGGAGAAGAAAGTCATTGTGTGGGAGCGGAGACAGCTGCTTGTGGGGACATT	3340
1001	LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp	1020
3341	TTACAGGACTTCCCGTGTCCGCCCGACTTGGTTCGGAGAGTCTCTCTTGGCCGAGCTGAT	3400
1021	GlyTyrThrSerLysGlyTrpSerLeuLeuAlaProIleThrAlaTyrrAlaGlnGlnThr	1040
3401	GGCTATACCTCCAAAGGGTGGAGTCTTCTCGCCCCCATCACTGCTTACGCCCAGCAGACA	3460
1041	ArgGlyLeuLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGluGlnAla	1060
3461	CGTGGCGCTTTTGGGCACCATAGTGTGTGAGCATGACGGGGCGGCACAAAGACAGAACAG	3520
1061	GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly	1080
3521	GGGGAATTCAGGTCTCTCCACAGTCATCAGTCTTCTCTCGGAACATCCATCTCGGG	3580
1081	ValLeuTrpThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro	1100
3581	GTTTTGTGGACTGTCTACCATGAGCTGGGCACAAAGACTCTTGSCCGGCTCACGGGGTCCG	3640
1101	ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTrpProSerProProGly	1120
3641	GTCAACGAGATGTACTTCAGTGTGAGGGGCACTTAGTAGGGTGGCCAGCCGCCCTGGG	3700

1121	ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArgAsn	1140
3701	ACTAAATCTTTGGAGCGGTGCACGTGTGGAGCGGTGCAGCTGTACTCGTCA	3760
1141	AlaAspValIleProAlaAAtgAtgArgGlyAspLysArgGlyAlaLeuLeuSerProArg	1160
3761	GCTGATGTATCCCGGCTCGAAGACGCGGGACAAACGGGGAGCGCTACTCTCCCGAGA	3820
1161	ProLeuSerThrLeuLysGlySerSerGlyGlyProValLeuCysProArgGlyHisAla	1180
3821	CCTCTTCCACCTTTGAAGGGTCTCTCAGAGCGCCGGTGTATGTCCTCCAGCGGGCCACGCT	3880
1181	ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle	1200
3881	GTCGGAGTCTTCCGGCGAGCTGTGTCTCTCGGGCGTGGCTAAAGTCCATAGATTTTCATC	3940
1201	ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro	1220
3941	CCCGTTGAGACACTCGACATCGTCACGCGGTCCCCACCTTTAGTGTACAACAGCACACCA	4000
1221	ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLys	1240
4001	CCTGCTGTGCCCCAGACCTATCAGGTCCGGTACTTGGCATGCCGCCACTGGCGAGTGGAAAG	4060
1241	SerThrLysValProValAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnPro	1260
4061	AGCACCAAAATTCCTGTGCGATATGCTGCTCAGGGGTATAAAGTGTCTAGTCTTAATCCC	4120
1261	SerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerLysAlaHisGlyIleLeuAsnPro	1280
4121	TCAGTGGGTGCCACCCCTCGGGTCTTGGGCGGTACTTGTCTAAAGGCACATGGCATCAATCCC	4180
1281	AsnIleArgThrGlyValArgThrValThrThrGlyAlaProIleThrTyrSerThrTyr	1300
4181	AACATTTAGACTTGGAGTCAGAGACTGTGACACCGGGGGCGCCCATCAGTACTTCCACATAT	4240
1301	GlyLysPheLeuAlaAspGlyGlyCysAlaGlyGlyAlaTyrAspIleIleIleCysAsp	1320
4241	GGCAAAATCTTCGCCGATGGGGCTGTGCGGCGCGGCTACGACATCATCATATGTGAT	4300
1321	GluCysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGlnAla	1340
4301	GAATGCCATCGGTGGATCTACACCACTTCTGGCATCGGAACAGTCTCTGATCAAGCA	4360
1341	GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProProGlySerValThr	1360
4361	GAGACAGCTGGGGTCAGACTAACTGTGTGGCTACAGCTACGCCCTCGGGTCAGTGACA	4420
1361	ThrProHisProAsnIleGluValAlaLeuGlyGlnGluGlyGluIleProPheTyr	1380
4421	ACCCCCACCCCAACATAGAGAGGTGGCCCTTGGGCAGGAGGGCGAGATCCCTCTTCTAT	4480
1381	GlyArgAlaIleProLeuSerTyrIleLysGlyValArgHisLeuIlePheCysHisSer	1400
4481	GGAGGGGGATTCCTCTGTCTTACATCAAGGGAGGAGACATCTGATCTCTTGGCATTTCA	4540
1401	LysLysLysCysAspGluLeuAlaAlaLeuArgGlyMetGlyLeuAsnSerValAla	1420
4541	AAGAAAAAGTGTGACGAGCTGCGCGCGGCCCTTCGGGGTATGGGCTTGAACCTCAGTGGCA	4600
1421	TyrTyrArgGlyLeuAspValSerValIleProThrGlnGlyAspValValValAla	1440
4601	TACTACAGAGGGTTCGAGCTCTCCCGTAATACCACTCAGGGAGACGTAGTGGTCTGCC	4660
1441	ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnVal	1460
4661	ACCGACGCCCTCATCACAGGGTATACTGGGGACTTTGACTCTCGTGTGACTGCAACGTA	4720
1461	AlaValThrGlnValValAspPheSerLeuAspProThrPheThrIleThrGlnIle	1480
4721	GGGGTCACTCAAGTTGTAGACTTCAGTTTAGACCCCAATTCACCAATAACCAACAGATT	4780

Qy	21	AspValLysPheProGlyGlyGlnIleValGlyValTyrLeuLeuProArgArg	40
Db	401	GAGCTTAAGTTTCGGCGCGCGCACATGTTGGCGAGTATATCTTTGTCGCGCAGG	460
Qy	41	GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly	60
Db	461	GGCCCCAGGTTGGTGTGCGCGCGCACAAAGAAAGATTCGGAGCGGTCCACAGCCACGTGGA	520
Qy	61	ArgArgGlnProIleProLysAspArgSerThrGlyLysSerThrGlyLysProGly	80
Db	521	AGCGCCAGCCCATCCCTAAAGATCGCGCTCCACTGGCAATCTCGGGAAACACAGGA	580
Qy	81	TyrProTrpProLeuTyrGlyAsnGlnGlyLeuGlyTrpAlaGlyTrpLeuLeuSerPro	100
Db	581	TACCCCTGGCCCTATACGGGAATGAGGAGCTCGGCTGGCGAGGATGGCTCTCTGTCGCC	640
Qy	101	ArgGlySerArgProSerTrpGlyProAsnAspProArgHisArgSerArgAsnValGly	120
Db	641	CGAGGTTCCCGTCCCTCTTGGGGCCCCAATGACCCCGGCATAGTTCGCGCAACGTGGGT	700
Qy	121	LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProValVal	140
Db	701	AGGTCACTCATACCTTAACGTGCGGCTTTGGCGACCTCATGGGGTACATCCCTGTGCTG	760
Qy	141	GlyAlaProLeuGlyGlyValAlaArgAlaLeuAlaHisGlyValArgValLeuGluAsp	160
Db	761	GGCGCCCGCTCGCGCGCGTCCGACAGCTCTCGCGCATGCGGTGAGAGTCTCTGGAGGAC	820
Qy	161	GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla	180
Db	821	GGGGTTAAATTTGCAACAGGGAATTAACCGGTTGCTCTCTTTCTCTTCTTCTTCTTCT	880
Qy	181	LeuLeuSerCysIleThrThrProValSerAlaAlaGluValLysAsnIleSerThrGly	200
Db	881	CTGCTGCTCCGATCACACCCCGTCTCGCTGCGCAAGTGAAGAACATCATACCTACCGC	940
Qy	201	TyrMetValThrAsnAspCysThrAsnAspSerIleThrTrpGlnLeuGlnAlaAlaVal	220
Db	941	TACATGCTGACTAACGACTGACCAATGACAGCATTTACCTGGCAGCTCCAGGCTGTGTC	1000
Qy	221	LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTrpIle	240
Db	1001	CTCCACGTCCTCCGGTGGTCCGTCGCGAGAAAGTGGGAATCATCTCAGTCTCGATA	1060
Qy	241	ProValSerProAsnValAlaValGlnArgProGlyAlaLeuThrGlnGlyLeuArgThr	260
Db	1061	CCGGTCTCACCGAATGGCGCGTGCAGCGCCCGCGCGCTCCACGAGGCTTGGCGAGC	1120
Qy	261	HisIleAspMetValValMetSerAlaThrLeuCysSerAlaLeuTyrValGlyAspLeu	280
Db	1121	CACATCGACATGTTGTGATGTCGCGCACGCTCTGCTCTGCTCTACGTTGGGGACCTC	1180
Qy	281	CysGlyGlyValMetLeuAlaAlaGlnMetPheIleValSerProGlnHisIleTrpPhe	300
Db	1181	TGCGGTGGGTGATGCTCGCAGCCCAATGTTTCAATGTCTCGCGCAGCACCACTGGTT	1240
Qy	301	ValGlnAspCysAsnCysSerIleTyrProGlyThrIleThrGlyHisArgMetAlaTrp	320
Db	1241	GTCCAGAGTCGAATGCTCCATCTACCTCGGTACCATCATCTGACACCGCATGGCATGG	1300
Qy	321	AspMetMetMetAsnTrpSerProThrAlaThrMetIleLeuAlaTyrAlaMetArgVal	340
Db	1301	GACATGATGATGAATGCTGCGCCACCGCTACCATGATCTTGGCGTACCGCATGCGTGC	1360
Qy	341	ProGluValIleIleAspIleIleSerGlyAlaHisIleTrpGlyValMetPheGlyLeuAla	360
Db	1361	CCCGAGTCAATATAGACATCAATAGCGGGCTCATTTGGGGCGTCAATGTTGGCTTGGCC	1420
Qy	361	TyrPheSerMetGlnGlyAlaTrpAlaLysValValIleLeuLeuAlaAlaGly	380
Db	1421	TACTTCTATGAGGAGCGTGGGGAAGTCTGTGTCATCTCTCTGTTGGCGCGCGG	1480
Qy	381	ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr	400
Db	1481	GTGGACCGCGCACCCATCTACTGTTGGGGTTCGCCGCGCAGACCAACCGCGCCCTCACC	1540
Qy	401	SerLeuPheAspMetClyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTrp	420
Db	1541	AGCTTATTTTGACATGGGCCCCAGGCAAAAATTCAGCTCGTTAAACCAATAGCAGCTGG	1600
Qy	421	HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer	440
Db	1601	CACATCAACCGCACCGCCCTGAACATGACTCTCTTGACACACCGGCTTTATCGGGTCT	1660
Qy	441	LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg	460
Db	1661	CTGTTCTACACCCACAGCTTCAACTCGTCAGGATGTCGGAACGCATGTCGCGCTGCCGC	1720
Qy	461	SerIleGluAlaPheArgValGlyTrpGlyAlaLeuGlnTyrGluAspAsnValThrAsn	480
Db	1721	AGTATCGAGGCTTCGGGTGGGATGGGCGGCTTTCGCAATATAGGATATATGTCACCAAT	1780
Qy	481	ProGluAspMetArgProTyrCysTrpHisTyrProProArgGlnCysGlyValValSer	500
Db	1781	CCAGAGGATATGAGACCTTATTTGCTGGCACTACCCACCAAGCAGTGTGGCTGTCTCC	1840
Qy	501	AlaLysThrValCysGlyProValTyrCysPheThrProSerProValValValGlyThr	520
Db	1841	GGCAAGACTGTGTGGCCAGTGTACTGTTTACCCCCAGCCAGTGTGTGGGCACG	1900
Qy	521	ThrAspArgLeuGlyAlaProThrTyrThrTrpGlyGluAsnGluThrAspValPheLeu	540
Db	1901	ACCGCAGGCTTTGGAGCGCCCACTTACAGCTGGGGGAGAAATGAGACAGATGCTTCTCTA	1960
Qy	541	LeuAsnSerThrArgProProLeuGlySerTrpPheGlyCysThrTrpMetAsnSerSer	560
Db	1961	TTGAACAGACCTCGACACCGCTGGGTCTGTTGGCTGACGCTGAGTGAATGAATCTTCT	2020
Qy	561	GlyTyrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer	580
Db	2021	GGCTACCAACAGACTTGGCGCGCACCCACCTCGCTACTAGAGTCTCAACGCCAGC	2080
Qy	581	ThrAspLeuLeuCysProThrAspCysPheArgLysHisProAspThrThrTyrLeuLys	600
Db	2081	ACGGACCTGTGTGGCCCCAGGACTGTTTTAGGAAGACTCTCTGATACCACTTACCTCAA	2140
Qy	601	CysGlySerGlyProTrpLeuThrProArgCysLeuLeuAspTyrProTyrArgLeuTrp	620
Db	2141	TGCGGCTCTGGGCGCTGGCTCAGCCAGGTGCTGATCGATCCTCCCTACAGGCTCTGG	2200
Qy	621	HisTyrProCysThrValAsnTyrThrIlePheLysIleArgMetTyrValGlyGlyVal	640
Db	2201	CATTACCCCTGCACAGTTAACTATACCATCTTCAAAATTAAGGATGTATGTGGAGGGTT	2260
Qy	641	GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp	660
Db	2261	GAGCAGGCTCACGGCTGCATGCAATTTTCACTCGTGGGGATCGTTGCAACTTGGAGGAC	2320
Qy	661	ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrGluTrpAlaIleLeuPro	680
Db	2321	AGAGACAGAGTCAACTGCTCTTGTGTGCACTCCACCGGAATGGGCCATTTTACCT	2380
Qy	681	CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuHisLeuHisGlnAsnIle	700
Db	2381	TGCTCTTACTCGGACCTGCCCTTGTGCACTGGTCTTCTCCACCTCCACCAAAACATC	2440
Qy	701	ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrLysTyrIleValArgTrp	720
Db	2441	GTGGAGGTACAATTCATGTATGGCTTATCACCTGGCTTACAAAATACATCGTCCGATGG	2500
Qy	721	GluTrpValIleLeuLeuPheLeuLeuAlaAspAlaArgValCysAlaCysLeuTrp	740
Db	2501	GAGTGGGTAAATACTCTTATCTCTGCTTGTAGCGGAGCGCAGGGTTGGCCCTGTATGG	2560
Qy	741	MetLeuIleLeuLeuGlyGlnAlaGluAlaLeuGluLysLeuValIleLeuHisAla	760

RESULT 2

AAC86937
ID AAC86937 standard; DNA; 9711 BP.

XX AAC86937;
AC AAC86937;

XX 02-APR-2001 (first entry)
DT

XX Nucleotide sequence of a hepatitis C virus (HCV) clone genotype 2a.
DE

XX Chimeric virus; bovine viral diarrhoea virus; BVDV; hepatitis C virus;
KW HCV; vaccine; viral inhibitor; antiviral; ss.

XX Hepatitis C virus.
OS

XX Location/Qualifiers
FH CDS

FT 341..942
/*tag= a

XX WO200075352-A2.
PN

XX 14-DEC-2000.
PD

XX 02-JUN-2000; 2000WO-US15527.
PF

XX 04-JUN-1999; 99US-0137817.
PR

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA

XX Nam J, Bukh J, Emerson SU, Purcell RH;
PI

XX WPI; 2001-071081/08.
DR

XX P-PSDB; AAB31168.
DR

XX New nucleic acid comprising a chimeric bovine viral diarrhoea virus
PT genome in which the (non-)structural region has been replaced by
PT hepatitis C virus (HCV) genome useful for treating or preventing HCV
PT signs and symptoms -

XX Disclosure; Page 82-85; 97pp; English.
PS

XX The specification describes a nucleic acid comprising a chimeric virus
CC genome, specifically bovine viral diarrhoea virus (BVDV) genome in which
CC the (non-) structural region has been replaced by the (non-) structural
CC region of a hepatitis C virus (HCV) genome. The nucleic acids comprising
CC the chimeric virus and the chimeric virus are useful for identifying
CC cell lines capable of supporting the replication of these chimeric
CC viruses, in screening for neutralizing antibodies to HCV of different
CC genotypes, in the production of HCV-BVDV virions, for the development
CC of inactivated or attenuated vaccines to prevent HCV-BVDV in a mammal,
CC in studying the molecular properties of HCV indirectly in vitro, and in
CC identifying inhibitors of viral enzyme activity which would be useful
CC as antiviral agents. Formulations or compositions comprising the
CC chimeric virions may be used to treat or prevent the signs and symptoms
CC of HCV. The present sequence represents a HCV clone, which is used
CC to construct chimeric nucleic acids of the invention.

XX Sequence 9711 BP; 1987 A; 2853 C; 2668 G; 2203 T; 0 other;
SQ

Alignment Scores:
Pred. No.:

Score: 16120.00 Length: 9711
Matches: 3033

Percent Similarity: 100.00% Conservativeness: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 22 Gaps: 0

US-09-980-559-2 (1-3033) x AAC86937 (1-9711)

Qy 1 MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgProGln 20

Db 341 ATGAGCACAATCTTAACCTCAAGAAACCAAAAGAAACCAACCTCGTCCCAAC 400

Qy	2681	IleTyrArgAlaCysSerLeuProGluGluAlaHisThrAlaIleHisSerLeuThrGlu	2700
Db	8381	ATATATCGGCGCTGCTCTTCCCGAGAGGAGCCACACATCCATACACTCGCTAACTGAG	8440
Qy	2701	ArgLeuTyrValGlyGlyProMetPheAsnSerLysGlyGlnThrCysGlyTyrArgArg	2720
Db	8441	AGACTTTACGTGGAGGGCCATGTTCAACAGCAAGGGCCAAACCTCGCGGTACAGCGT	8500
Qy	2721	CysArgAlaSerGlyValLeuThrThrSerMetGlyAsnThrIleThrCysTyrValLys	2740
Db	8501	TGCGCGCCAGCGGGTCTCACCACATAGCATGGGGAACACCATCATCTCTACGTCAA	8560
Qy	2741	AlaLeuAlaAlaCysValAlaAlaGlyIleIleAlaProThrMetLeuValCysGlyAsp	2760
Db	8561	GCCTTAGCGGCTGTAAAGCTGCAGGATTAATCGCCCAATGCTGGTATCGCGCAT	8620
Qy	2761	AspLeuValValIleSerGluSerGlnGlyThrGluGluAspGluArgAsnLeuAla	2780
Db	8621	GACTTGGTTGTCATCTCAGAAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	8680
Qy	2781	PheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProArgProGluTyr	2800
Db	8681	TTCAAGGAGGCTATGACAGGATATCTGCGCCCTCTCTGTCGACCCCGCCAGCGGAT	8740
Qy	2801	AspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaLeuGlyProGlnGly	2820
Db	8741	GATCTGGAGCTGATAAATCTTGTCTCTCAATGTGTGTGGCGCTGGGCCCAAGGC	8800
Qy	2821	ArgArgArgTyrTyrLeuThrArgAspProThrProIleAlaArgAlaAlaTyrGlu	2840
Db	8801	CGCCGAGATACCTACCTGACGAGAGCCCTACCACTCCATCGCCCGGCTGCTGGAA	8860
Qy	2841	ThrValArgHisSerProValAsnSerTrpLeuGlyAsnIleIleGlnTyrAlaProThr	2860
Db	8861	ACAGTAGACACTCCCTCTCAATTCATGCTGGGAAACATCATCTAGTACGCCCGACC	8920
Qy	2861	IleTyrAlaArgMetValLeuMetThrHisPhePheSerIleLeuMetAlaGlnAspThr	2880
Db	8921	ATATGGGCTCGCATGGTCTGATACACACTTCTTCTCCATCTCATGGCTCAAGACG	8980
Qy	2881	LeuAspGlnAsnLeuAsnPheGluMetTyrGlyAlaValTyrSerValSerProLeuAsp	2900
Db	8981	CTGACACAGAACTCACTTTCAGATGATACGAGCGGTACTCCGTGAGTCCCTTGAC	9040
Qy	2901	LeuProAlaIleLeuGluArgLeuHisGlyLeuAspAlaPheSerLeuHisThrTyrThr	2920
Db	9041	CTCCGAGCTATAATTTGAAAGGTTACATGGGCTTGACGCTTTCTCTGCAACATAC	9100
Qy	2921	ProHisGluLeuThrArgValAlaSerAlaLeuArgLysLeuGlyAlaProProLeuArg	2940
Db	9101	CCCCAGCAACTGACACGGTGGCTTCAGCCCTCAGAAACTTGGGGCGCCACCCCTCAG	9160
Qy	2941	AlaTrpLysSerArgAlaArgAlaValArgAlaSerLeuIleSerArgGlyGlyArgAla	2960
Db	9161	CGGTGGAGAGAGCCGGCAGTGCAGTGCAGGCGTCCCTCATCTCCCGTGGGGGAGCG	9220
Qy	2961	AlaValCysGlyArgTyrLeuPheAsnTrpAlaValLysThrLysLeuLysLeuThrPro	2980
Db	9221	GCCGTTTTCGGTTCGATATCTCTTCAATTTGGGGGCTGAAGCAAGCTCAAACTCACT	9280
Qy	2981	LeuProGluAlaArgLeuLeuAspLeuSerSerTrpPheThrValGlyAlaGlyGlyGly	3000
Db	9281	TTGCGGAAGCGCGCTCTCGATTTATCCAGCTGGTTTACCGTGGCGCGCGGGGGG	9340
Qy	3001	AspIleTyrHisSerValSerArgAlaArgProArgLeuLeuLeuPheGlyLeuLeuLeu	3020
Db	9341	GACATTTATCAGGGTGTGGGTGCGGAGCCCGGCTTATTTGCTTTGGCTTACTCTTA	9400
Qy	3021	LeuPheValGlyValGlyLeuPheLeuLeuProAlaArg 3033	
Db	9401	CTTTTGTAGGGTAGGCTTTTCTCTACTCCCGCGCTCGG 9439	

Qy	1221	ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLys	1240	Qy	1581	LysGlnSerGlyGluAsnPheAlaTyrLeuThrAlaTyrGlnAlaThrValCysAlaArg	1600
Db	4001	CTGCTGTGCCCGACGCTATACGTCGGGTACTTGCATGCCGCCGACTGGCAGTGGAAAG	4060	Db	5081	AAGCAATCGGGGAAAAATTTGGCACTTAAACAGCTACAGCTACAGTGTGCGCTAGG	5140
Qy	1241	SerThrLysValProValAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnPro	1260	Qy	1601	AlaLysAlaProProProSerTrpAspValMetTTrpLysCysLeuThrArgLeuLysPro	1620
Db	4061	AGCACCAAGTCTCTGTGCGATATGCTGCTCAGGGGTATAAAGTGTAGTGTCTTAATCCC	4120	Db	5141	GCCAAAGCCCCCCCCCGCTCCGGACGCTCATGTGGAAGTGTTCGACTCGCATCAAGCCC	5200
Qy	1261	SerValAlaThrLeuGlyPheGlyAlaTyrLeuSerLysAlaHisGlyLysAsnPro	1280	Qy	1621	ThrLeuValGlyProThrProLeuLeuTyrArgLeuGlySerValThrAsnGluValThr	1640
Db	4121	TCAGTGGCTGCCACCTCGGGGTTTGGGGCTACTTGTCTAAGGCACATGTCATCAATCCC	4180	Db	5201	ACACTCGTGGGCCCCACACCTCTCTGTACCGCTTGGGCTCTGTTCACCAACGAGGTCAAC	5260
Qy	1281	AsnIleArgThrGlyValArgThrValThrGlyAlaProIleThrTyrSerThrTyr	1300	Qy	1641	LeuThrHisProValThrLysTyrIleAlaThrCysMetGlnAlaAspLeuGluValMet	1660
Db	4181	AACATTAGGACTGGAGTCAGACTGTGACACCGGGCGCCCATCAGTACTCCACATAT	4240	Db	5261	CTCACACATCCCGTCGACGAAATACATCGCCACCTGCATGCAAGCCGACCTTGAGGTCAATG	5320
Qy	1301	GlyLysPheLeuAlaAspGlyGlyCysAlaGlyAlaTyrAspIleIleCysAsp	1320	Qy	1661	ThrSerThrTrpValLeuAlaGlyValLeuAlaAlaValAlaAlaTyrCysLeuAla	1680
Db	4241	GGCAAAATTCCTCGCGATGGGGCTGTGGCGGGCGGCTTACGACATCATATATGTAT	4300	Db	5321	ACCAGCACATGGGTCTTGGCAGGGGAGTCTTGGGGCGCGCTCGCCGCTATTGCGCTGGCG	5380
Qy	1321	GluCysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGlnAla	1340	Qy	1681	ThrGlyCysValCysIleIleGlyArgLeuHisIleAsnGlnArgAlaValValAlaPro	1700
Db	4301	GAATGCCATCGCGTGGTCTTACACCATCTTGGCATCGGAACAGTCTTGTATCAAGCA	4360	Db	5381	ACCGGGTGTGTGTGCATCATCGCGCTTGCACATTAACGAGGAGCGCTCGTTGCGCGG	5440
Qy	1341	GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProProGlySerValThr	1360	Qy	1701	AspLysGluValLeuTyrGluAlaPheAspGluMetGluGluCysAlaSerArgAlaAla	1720
Db	4361	GAGACACTGGGTGACACTTAACGTGCTGTGCTACAGTACGCCCCCTGGGTGAGTGACA	4420	Db	5441	GACAGGAGGTCTCTATGAGGCTTTTGTAGATGAGAGGAATGTGCTCTTAGGGCGGCT	5500
Qy	1361	ThrProHisProAsnIleGluGluValAlaLeuGlyGlnGlyGluIleProPheTyr	1380	Qy	1721	LeuIleGluGlyGlnArgIleAlaGluMetLeuLysSerLysIleGlnGlyLeuLeu	1740
Db	4421	ACCCCCACCCACATAGAGGAGTGGCCCTTGGGCGAGGCGGAGATCCCTTCTAT	4480	Db	5501	CTCATTTGAAGAGGGGAGGATAGCCGAGATGCTGAAGTCCAAAGATCCAGGCTTATTG	5560
Qy	1381	GlyArgAlaIleProLeuSerTyrIleLysGlyGlyArgHisIleIlePheCysHisSer	1400	Qy	1741	GlnGlnAlaSerLysGlnAlaGlnAspIleGlnProThrValGlnAlaSerTrpProLys	1760
Db	4481	GGGAGGGCGATTCCCTCTCTTACATCAAGGAGGAAAGACATCTGATCTTCTGCCCATCA	4540	Db	5561	CAGCAAGCTTCCAAACAGCTCAAGACATCAACCCACTGTGCGAGGCTTCATGGCCCAAG	5620
Qy	1401	LysLysLysCysAspGluLeuAlaAlaAlaLeuArgGlyMetGlyLeuAsnSerValAla	1420	Qy	1761	ValGluGlnPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeuAla	1780
Db	4541	AAGAAAAGTGTGACGAGCTCGCGCGCGCCCTTGGGGTATGGGCTTGAACCTAGTGCGCA	4600	Db	5621	GTAGAACAATTTCTGGGCCAAACACATGTGGAATCTTATTAGCGGCATCCAAATACCTCGCA	5680
Qy	1421	TyrTyrArgGlyLeuAspValSerValIleProThrGlnGlyAspValValValAla	1440	Qy	1781	GlyLeuSerThrLeuProGlyAsnProAlaValAlaSerMetMetAlaPheSerAlaAla	1800
Db	4601	TACTACAGAGGTTGGAGCTCTCGTAATACCAACTCAGGAGAGGAGTGTGCTGCC	4660	Db	5681	GGACTATCAACACTGCGCAGGAAACCTGCGAGTGTTCATGATGCGGTTCAGTGGCGCC	5740
Qy	1441	ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnVal	1460	Qy	1801	LeuThrSerProLeuSerThrSerThrThrIleLeuLeuAsnIleLeuGlyTyrTrpLeu	1820
Db	4661	ACCGACGCCCTCATGACAGGGTATCTGCGGACCTTGTACTCCGTGATCGACTGCAACGTA	4720	Db	5741	CTCACGAGTCCGCTGTCAACAGCACCATCTCTCTCAACATTTTGGGGGCTGGCTA	5800
Qy	1461	AlaValThrGlnValValAspPheSerLeuAspProThrPheThrIleThrThrGlnIle	1480	Qy	1821	AlaSerGlnIleAlaProProAlaGlyAlaThrGlyPheValValSerGlyLeuValGly	1840
Db	4721	CGCGTCACTCAAGTTGTAGACTTTCAGTTTAGCCCCACATTCACATCAACACACAGATT	4780	Db	5801	GCATCCCAATTTGCACACCCCGCGGGGCCACTGGCTTCGTTCAGTGGCTTAGTGGGA	5860
Qy	1481	ValProGlnAspAlaValSerArgSerGlnArgArgGlyArgThrGlyArgGlyArgLeu	1500	Qy	1841	AlaAlaValGlySerIleGlyLeuGlyLysValLeuValAspIleLeuAlaGlyTyrGly	1860
Db	4781	GTCCCTCAAGACGCTGTCTCAGTAGCCAGCGCGGGGTGCAAGGTAGGGGAAGACTG	4840	Db	5861	GCTGCGGTAGGAGTATAGGCTTAGTAAAGTGTCTAGTGACATCTCTGGCAGGAGTATGGT	5920
Qy	1501	GlyIleTyrArgTyrValSerThrGlyGluArgAlaSerGlyMetPheAspSerValVal	1520	Qy	1861	AlaGlyIleSerGlyAlaLeuValAlaPheLysIleMetSerGlyGlyLysProSerMet	1880
Db	4841	GGCATTTATAGTATGTTTTCACCTGGTGAGCGAGCCTCAGGAATGTTTTCAGAGTGTAGT	4900	Db	5921	CGCGGCAATTTCCGGGCTCTCTGTCGATTCAGATCATGTCTGCGGAGAGCCCTCCATG	5980
Qy	1521	LeuCysGluCysTyrAspAlaGlyAlaAlaTrpTyrGluLeuThrProSerGluThrThr	1540	Qy	1881	GluAspValValAsnLeuLeuProGlyIleLeuSerProGlyAlaLeuValValGlyVal	1900
Db	4901	CTCTGTAGTGTCTACGACGCGAGGGCCGATGTTAGTATGACTCACCCATCGAGACCCACC	4960	Db	5981	GAGGATGCTGCTCAACTTGTCTGCTGGAATTCGTCTCCGGGTGCTTGGTAGTGGAGTC	6040
Qy	1541	ValArgLeuArgAlaTyrPheAsnThrProGlyLeuProValCysGlnAspHisLeuGlu	1560	Qy	1901	IleCysAlaAlaIleLeuArgArgHisValGlyProGlyGlyAlaValGlnTrpMet	1920
Db	4961	GTGAGGCTCAGGGCGTATTTCAACACGCGCGGTTTGGCTGTGTGCCAAGACCATCTTGAG	5020	Db	6041	ATCTGCGGGCCCAATTTGCGCCGACAGTGGGACCGGGGAGGCGCGCTCCAAATGATG	6100
Qy	1561	PheTrpGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThr	1580	Qy	1921	AsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaProThrHisTyrValThr	1940
Db	5021	TTTTGGGAGGCGATTTTCAACCGGCTCACACACATAGATGCCCATCTTCTTTCCCAACA	5080	Db	6101	AATAGACTCATGCTTTGCTTCCAGAGGAATACAGTCCGCCCCCACTAGTGTGACG	6160
				Qy	1941	GluSerAspAlaSerGlnArgValThrGlnLeuLeuGlySerLeuThrIleThrSerLeu	1960

Db	1781		CCAGAGGATGAGACCCCTTATGCTGGCACTACCCACCAAGGAGGTGGTGGTCTCC	1840
Qy	501		AlaLysThrValCysGlyProValTyrCysPheThrProSerProValValGlyThr	520
Db	1841		GCAGAGCTGTGTGGCCAGTGACTGTTCACCCCAAGCCAGTGGTAGTGGGACG	1900
Qy	521		ThrAspArgLeuGlyAlaProThrTyrThrTyrGlyGluAsnGluThrAspValPheLeu	540
Db	1901		ACCGACAGGCTTGGAGGCCCACTTACACGTGGGGGAGAAATGAGACAGATGCTTCCTA	1960
Qy	541		LeuAsnSerThrArgProProLeuGlySerTyrPheGlyCysThrTyrMetAsnSerSer	560
Db	1961		TTCAACAGCACTCGACCAACCGCTGGGTCTATGCTGGCTGCACGTGGTGAACCTCTCT	2020
Qy	561		GlyTyrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer	580
Db	2021		GGCTACACCAAGACTTGGCGCGCACCAACCTGCGGTACTAGAGCTCACTTCAACGCCAGC	2080
Qy	581		ThrAspLeuLeuCysProThrAspCysPheArgLysHisProAspThrTyrThrLeuLys	600
Db	2081		ACGGACCTGTGTGCCCCAGGACTGTTTAGGAAGCATCTGATACCACTTACCTCAAA	2140
Qy	601		CysGlySerGlyProThrLeuThrProArgCysLeuLeuAspTyrProTyrArgLeuTyr	620
Db	2141		TGGGGCTCTGGGCCCTGGCTCAGCCCAAGGTGCTGATCGACTACCCCTACAGGCTCTGG	2200
Qy	621		HisTyrProCysThrValAsnTyrThrIlePheLysIleArgMetTyrValGlyVal	640
Db	2201		CATTACCCCTGGCACAGTTAACTATACCATCTTCAAAATAGGATGTATGTGGAGGGGT	2260
Qy	641		GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp	660
Db	2261		GAGCACAGGCTCAGGCTGATGCAATTCACCTCGTGGGATCGTTGCAACTTGGAGGAC	2320
Qy	661		ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGluTyrPheAlaLeuPro	680
Db	2321		AGAGACAGAAGTCAACTGCTCTCTGTTGTGCACTCCACACCGAATGGGCCATTTTACCT	2380
Qy	681		CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisLeuHisGlnAsnIle	700
Db	2381		TGCTCTTACTCGGACCTGCCCGCTTGTGACATGGTCTTCTCCACCTCCACCAAAACATC	2440
Qy	701		ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrLysTyrIleValArgTyr	720
Db	2441		GTGGACGTACAAATTCATGATGCGCTATCACTTGCCTTCAAAATACATGCTCCGATGG	2500
Qy	721		GluTyrValIleLeuLeuPheLeuLeuLeuAlaAspAlaArgValCysAlaCysLeuTyr	740
Db	2501		GAGTGGGTAATACTTATTCTCTCTTAGCGGACGCCAGGCTTTGCGCTGTCTTATGG	2560
Qy	741		MetLeuIleLeuLeuGlyGlnAlaGluAlaAlaLeuGluLysLeuValIleLeuHisAla	760
Db	2561		ATGCTCATCTTGTGGCCAGGCGGAGGAGGAGCACTAGAGAAGCTGTGCTATCTTTCACGCT	2620
Qy	761		AlaSerAlaAlaSerCysAsnGlyPheLeuTyrPheValIlePhePheValAlaAlaTyr	780
Db	2621		CGAGGCGCAGCTAGCTGCAATGCTTCCATATTTTGTATTTTGTGCTGCTGCTGG	2680
Qy	781		TyrIleLysGlyArgValProLeuAlaThrTyrSerLeuThrGlyLeuTyrSerPhe	800
Db	2681		TACATCAAGGGTCGGGTAGTCCCTTAGCTTACCTATTTCCCTCACTGGCTGTGCTCTT	2740
Qy	801		SerLeuLeuLeuLeuAlaLeuProGlnGlnAlaTyrAlaTyrAspAlaSerValHisGly	820
Db	2741		AGCCTACTGCTCTAGCATTTGCCCAACAGGCTTATGCTTATGACGCATCTGTGCAATGGC	2800
Qy	821		GlnIleGlyAlaAlaLeuLeuValMetIleThrLeuPheThrLeuThrProGlyTyrLys	840
Db	2801		CAGATAGGAGCGGCTCTGCTGGTAAATGATCACTCTCTTACTCTCACCCCGGGTATAAG	2860
Qy	841		ThrLeuLeuSerArgPheLeuTyrTyrProLeuCysTyrLeuLeuThrLeuGlyAlaMet	860
Db				
Db	2861		ACCTTCTCAGCGGCTTTTGTGGTGTGTGCTATCTTCTCAGCCCTGGGGAGAGCTATG	2920
Qy	861		ValGlnGluTyrAlaProProMetGlnValArgGlyGlyArgAspGlyIleIleTyrAla	880
Db	2921		GTCCAGGAGTGGGCAACCACTATGAGGTGCGGTGGCGGTGATGGCATATATGGGCC	2980
Qy	881		ValAlaIlePheTyrProGlyValValPheAspIleThrLysThrLeuLeuAlaValLeu	900
Db	2981		GTCCGCAATATTACCCAGGTGTGGTGTGTGACATAACCAAGTGGCTCTTGGCGGTCTT	3040
Qy	901		GlyProAlaTyrLeuLeuLysGlyAlaLeuThrArgValProTyrPheValArgAlaHis	920
Db	3041		GGGCTCTGTCTTACCTCTAAAGGTGCTTTGACGCGCGTCCCTACTTCTGTCAGGGCTCAC	3100
Qy	921		AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyGlyArgTyrValGlnMet	940
Db	3101		GCTCTACTGAGATGTGCACCATGCAAGGCATCTCGCGGGGGGAGGTACGTCCAGATG	3160
Qy	941		AlaLeuLeuAlaLeuGlyArgTyrThrGlyThrTyrIleTyrAspHisLeuThrProMet	960
Db	3161		GGCTACTAGCCCTTGGCAGGTGGACTGGCACTTACATCTATGACCACTCACCCCTATG	3220
Qy	961		SerAspTyrAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIleIlePhe	980
Db	3221		TCGGATTTGGGCTGTAGTGGCTTGGGGACCTGGCGGTGCGGTGTGAGCTATCATCTTC	3280
Qy	981		SerProMetGluLysLysValIleValTyrGlyAlaGluThrAlaAlaCysGlyAspIle	1000
Db	3281		AGTCGATGGAGAAAGTCAITGTGGGGAGCGGAGACAGCTGCTTGTGGGGACATT	3340
Qy	1001		LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp	1020
Db	3341		TTACACGAGACTTCCGCTGTCGCCCGACTTGTGTGGAGGTCTCTCTTGGCCAGCTGAT	3400
Qy	1021		GlyTyrThrSerLysGlyTyrSerLeuLeuAlaProIleThrAlaTyrAlaGlnThr	1040
Db	3401		GGCTATACCTCCAGGGGTGGAGTCTTCTCGCCCATCACCTGCTTACGCCAGCAGACA	3460
Qy	1041		ArgGlyLeuLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGluGlnAla	1060
Db	3461		CGTGCCCTTTGGGCACCATAGTGTGAGCATGAGCGGGCGGCAAGACAGAACAGGCT	3520
Qy	1061		GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly	1080
Db	3521		GGGGAATTCAGTCTCTGTCCACAGTCACTCACTCTCTCGGAACATCATCTCGGG	3580
Qy	1081		ValLeuTyrThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro	1100
Db	3581		GTTTTGTGAGCTGTCTACCATGGAGCTGGCAACAAAGACTCTGGCGGCTCAGCGGTCG	3640
Qy	1101		ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTyrProSerProGly	1120
Db	3641		GTCCAGCAGATGATCTCCAGTGTGAGGGGACTTAGTAGGGTGGCCAGCCCTCGG	3700
Qy	1121		ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArgAsn	1140
Db	3701		ACTAATCTTTGGAGCCGTGCACGTGTGGAGCGGTGACCTGTACTTGTTCACGCGAAC	3760
Qy	1141		AlaAspValIleProAlaArgArgGlyAspLysArgGlyAlaLeuLeuSerProArg	1160
Db	3761		GCTGATGTCTCCGGCTCGAAGACCGGGGCAACAAACGGGAGCGCTACTCTCCCGAGA	3820
Qy	1161		ProLeuSerThrLeuLysGlySerSerGlyGlyProValLeuCysProArgGlyHisAla	1180
Db	3821		CCTCTTCCACTTGAAGGGTCTCAGAGGCGCGGTGTATGCCCCAGGGGCGCACGCT	3880
Qy	1181		ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle	1200
Db	3881		GTCCGAGTCTTCCGGGAGCTGTGTCTCTCGGGCGGTGGCTAAGTCCATAGATTTCATC	3940
Qy	1201		ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro	1220
Db	3941		CCCGTTGAGACACTCGACATCGTCAGCGGTCCCCCACCTTTAGTGACAAACAGCACCA	4000

FT XX /product= "Hepatitis C virus protein"

PN XX HQ200075338-A2.

PD XX 14-DEC-2000.

XX XX 02-JUN-2000; 2000WO-US15446.

XX PF 04-JUN-1999; 99US-0137693.

XX PR (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PA Yanagi M, Bukh J, Emerson SU, Purcell RH;

XX PI WPI; 2001-061728/07.

XX DR P-PSDB; AAB30729.

XX PT Nucleic acid molecule encoding human hepatitis C virus of genotype 2a
for developing vaccines, for diagnosis of hepatitis C virus and in
screening assays for identification of antiviral agents -

XX PS Claim 3; Page 84-87; 167pp; English.

XX CC The present sequence represents a nucleotide sequence of infectious
Hepatitis C virus (HCV) strain HC-J6CH genotype 2a. The HCV
polynucleotide sequence is capable of expressing the virus when
transfected into cells. The HCV protein is useful for assaying candidate
antiviral agents for activity against HCV. Antibodies specific for HCV
polypeptide are useful in prevention and treatment of diseases caused
by HCV in animals, in particular humans. The HCV polypeptides serve as
immunogens in the development of vaccines for preventing HCV in mammals
or as antigens in diagnostic assays for detecting the presence of HCV
in biological samples. The HCV polynucleotide is also useful for
identifying cell lines capable of supporting the replication of HCV
in vitro and to produce attenuated viral strains via passage in vitro
or in vivo.

XX SQ Sequence 9711 BP; 1987 A; 2853 C; 2668 G; 2203 T; 0 other;

Alignment Scores:

Pred. No.: 0 Length: 9711
Score: 16120.00 Matches: 3033
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

US-09-980-559-2 (1-3033) x AAC86644 (1-9711)

QY 1 MetSerThrAnProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGln 20
DB 341 ATGAGCAAAATCTTAAACCTCAAGAAACCAAGAAACCAACCAACCTGCGCCCAAA 400
QY 21 AspValLysPheProGlyGlyGlyGlnLeuValGlyGlyValTyrLeuLeuProArgArg 40
DB 401 GACGTTAAGTTTCCGGCGCGCGCCAGATCGTTGCGGAGTATACTTGTTCGCGCAGG 460
QY 41 GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly 60
DB 461 GGCCCGCAGGTTGGGTGTGCGCGCGCACCAAGGAAGACTTCGGAGCGGTCCCGACCGTGA 520
QY 61 ArgArgGlnProLysPheProLysAspArgSerThrGlyLysSerTyrGlyLysProGly 80
DB 521 AGGCGCCAGCCCAATCCCTTAAAGATCGCGCTCCACTGGCAATCTTGGGGAACCAAGGA 580
QY 81 TyrProTTPProLeuTyrGlyAsnGluGlyLeuGlyTyrAlaGlyTyrLeuLeuSerPro 100
DB 581 TACCCCTGCGCCCTATACGGGAATGAGGACTTCGCTGGGCGAGATGCTCTGTCCCCC 640
QY 101 ArgGlySerArgProSerTyrGlyProAsnAspProArgHisArgSerArgAsnValGly 120
DB 641 CGAGGTTCCCGTCCCTCTTGGGGGCCCAATGACCCCGGCATAGGTGCGCGCAACGTGGT 700

QY 121 LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProValVal 140
DB 701 AAGGTCATCGATACCTAAACGTGCGGCTTTCGCGACCTCATGGGCTACATCCCTGTCGTG 760
QY 141 GlyAlaProLeuGlyGlyValAlaAlaArgAlaLeuAlaHisGlyValArgValLeuGluA 160
DB 761 GCGCGCCCGCTCGCGCGGCTGCCAGAGCTCTCGCGCATGGCGTGGAGTCTCTGGAGGAC 820
QY 161 GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla 180
DB 821 GGGGTTAATTTTGCACAGGGAACCTTACCGGTGCTCTCTTCTATCTCTTCTGCTGGCC 880
QY 181 LeuLeuSerCysIleThrThrProValSerAlaAlaGluValValAsnIleSerThrGly 200
DB 881 CTGCTGTCTGCATACCAACCGCGCTCTCGCGTGCAGAGTGAAGAACATCATAGTACCGGC 940
QY 201 TyrMetValThrAsnAspCysThrAsnAspSerIleThrTyrGlnLeuGlnAlaVal 220
DB 941 TACATGGTGACTTAACGACTGCACCAATGACACATTTACCTGGCAGCTCCAGGCTGCTGTC 1000
QY 221 LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTyrIle 240
DB 1001 CTCACGCTCCCGGCTGCTCGCTCGGAGAAAGTGGGGAATGCATCTCAGTGTCTGATA 1060
QY 241 ProValSerProAsnValAlaValGlnArgProGlyAlaLeuThrGlnGlyLeuArgThr 260
DB 1061 CCGGCTCTACCGAAATGTGCGCTGCAGCGCGCGCGCTCAGCGAGGCTTTCGCGAGC 1120
QY 261 HisIleAspMetValValMetSerAlaThrLeuCysSerAlaLeuTyrValGlyAspLeu 280
DB 1121 CACATCGCATGGTGTGTATGTCGCGCACGCTGCTCTGCGCTCTACCTGGGGACCTTC 1180
QY 281 CysGlyGlyValMetLeuAlaAlaGlnMetPheIleValSerProGlnHisThrPhe 300
DB 1181 TCGCGTGGGGTGTATGTCGACGCCCAATGTTCAATGTTCTCGCGCGCAGCACCATCGTTT 1240
QY 301 ValGlnAspCysAsnCysSerIleTyrProGlyThrIleThrGlyHisArgMetAlaTyr 320
DB 1241 GTCCAAAGACTGCAATTTGCTCCATCTACCTCGGTGTACCATCACTGACACCGCATGG 1300
QY 321 AspMetMetMetAsnTrpSerProThrAlaThrMetIleLeuAlaTyrAlaMetArgVal 340
DB 1301 GACATGATGATGAATCGGTGCGCCACCGCTACCATGATCTTGGCGTACCGATGCGGTGTC 1360
QY 341 ProGluValIleIleAspIleIleSerGlyAlaHisTyrGlyValMetPheGlyLeuAla 360
DB 1361 CCGAGGTCTATTATACATCATTAGCGGGGCTCATTTGGGCGCTCATGTTGCGCTGGCC 1420
QY 361 TyrPheSerMetGlnGlyAlaTyrAlaLysValValIleLeuLeuLeuAlaAlaGly 380
DB 1421 TACTTCTCTATGCGAGGCGGTGGCGCAAAAGTCGTTGTCTATCTCTGTTGGCGCGCGG 1480
QY 381 ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr 400
DB 1481 GTGGAGCGCGCACCATCATCTGTTGGGGTTCGCGCGCGCAGACACCGCGCGCTCACC 1540
QY 401 SerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTyr 420
DB 1541 AGCTTTATTTGACATGGCGCCCGCAGGAGAAATCCAGCTCGTTAACCAACCAATGGCAGCTG 1600
QY 421 HisIleAsnArgThrAlaLeuAsnCysAsnAspSerIleHisThrGlyPheIleAlaSer 440
DB 1601 CACATCAACCGCACCGCTGAACTGCAATGACTCTCTTGCACACCGGCTTTATCGCGTCT 1660
QY 441 LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg 460
DB 1661 CTGTTCTTACACCGCACAGCTTCAACTCGTCAGGATGTCGCGAAGCATGTCCGCTGCGCTG 1720
QY 461 SerIleGluAlaPheArgValGlyTyrGlyValAlaLeuGlnTyrGluAspAsnValThrAsn 480
DB 1721 AGTATCGAGGCTTTCGGGTGGATGGGGCGCTTGCATATGAGGATATATGTCACCAAT 1780
QY 481 ProGluAspMetArgProTyrCysTyrHisTyrProProArgGlnCysGlyValValSer 500

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 11, 2004, 23:08:29 ; Search time 1476 Seconds
(without alignments)
5547.016 Million cell updates/sec

Title: US-09-980-559-2
Perfect score: 16120
Sequence: 1 MSTNPKPQKTKNTNRRPQ.....LLFGLLLVGVGLFLPAR 3033

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DBV=xlh
-Q=/cgn2_1/USPTO.spool/US0980559/runat_09022004_141953_26393/app_query.fasta_1.3207
-DB=N Geneseq 19Jun03 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=bloum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPWT=ptp -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0980559 @cgn_1_1082 @runat_09022004_141953_26393 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq 19Jun03:
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	16120	100.0	9711	22	AAC86644	Nucleotide sequenc
2	16120	100.0	9711	22	AAC86937	Nucleotide sequenc
3	16106	99.9	9711	22	AAF23486	Infectious Hepatit
4	15824	98.2	9589	14	AAQ38218	NANBH virus strain
5	14943	92.7	9678	24	ABK88904	Human HCV-related
6	13810	85.7	9511	14	AAQ38221	NANBH virus strain
7	13049	80.9	9611	22	AAC86645	Nucleotide sequenc
8	13049	80.9	9611	22	AAC86647	Nucleotide sequenc
9	12925	80.2	9611	22	AAC86646	Nucleotide sequenc
10	12925	80.2	9611	22	AAC86648	Nucleotide sequenc
11	12034.5	74.7	9605	24	ABK91431	Hepatitis C virus
12	12033.5	74.6	9605	24	ABK91428	Hepatitis C virus
13	12031.5	74.6	9605	24	ABK91411	Hepatitis C virus
14	12031.5	74.6	9605	24	ABK91429	Hepatitis C virus
15	12028.5	74.6	9605	24	ABK91425	Hepatitis C virus
16	12027.5	74.6	9605	24	ABK91426	Hepatitis C virus
17	12027.5	74.6	9605	24	ABK91432	Hepatitis C virus
18	12027.5	74.6	11076	21	AAA98965	Hepatitis C virus
19	12026.5	74.6	9605	24	ABK91433	Hepatitis C virus
20	12025.5	74.6	9605	24	ABK91424	Hepatitis C virus
21	12025.5	74.6	9605	24	ABK91430	Hepatitis C virus
22	12025.5	74.6	9605	24	AAAD25332	Hepatitis C virus
23	12025.5	74.6	11062	24	AAAD25331	Hepatitis C virus
24	12022	74.6	9608	24	ABK91427	Hepatitis C virus
25	11998.5	74.4	9413	25	AAAD25517	Hepatitis C virus
26	11998.5	74.4	9413	25	AAAD36555	Hepatitis C virus
27	11992.5	74.4	9413	25	AAAL53723	Hepatitis C virus
28	11990.5	74.4	9413	16	AAQ81559	Hepatitis C virus
29	11989.5	74.4	9416	20	AAQ59394	Non-A, non-B hepat
30	11986.5	74.4	9413	16	AAQ03960	Partial HCV non-st
31	11984.5	74.3	9416	13	AAQ21829	Non-A, non-B viral
32	11984.5	74.3	9416	13	AAQ20268	Non-A, non-B viral
33	11973	74.3	9646	19	AAV59361	Hepatitis C virus
34	11973	74.3	9646	24	ABK87285	cDNA encoding hepa
35	11973	74.3	12980	19	AAV59364	Hepatitis C virus
36	11973	74.3	12980	24	ABK87286	Hepatitis C virus
37	11966	74.2	10803	25	ABX10617	MKO-Z nucleotide s
38	11964	74.2	9599	20	AAQ24832	Infectious hepatit
39	11964	74.2	9599	22	AAC86938	Nucleotide sequenc
40	11963.5	74.2	9413	16	AAQ80498	DNA encoding HCV p
41	11961	74.2	9599	22	AAF23491	Infectious Hepatit
42	11958.5	74.2	9595	20	AAQ24843	Infectious Hepatit
43	11958.5	74.2	9595	22	AAC86939	Nucleotide sequenc
44	11955.5	74.2	9595	22	AAF23492	Infectious Hepatit
45	11945	74.1	9487	16	AAQ79140	Hepatitis C virus

ALIGNMENTS

RESULT 1
AAC86644
ID AAC86644 standard; DNA; 9711 BP.
XX
AC AAC86644;
XX
DT 02-APR-2001 (first entry)
XX
DE Nucleotide sequence of infectious Hepatitis C virus strain HC-J6CH.
XX
KW HCV; HCV strain HC-J6CH; HCV genotype 2a; antiviral; vaccine; ss.
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT CDS 341..9442
/*tag= a

Db 8167 ACACAAAAGCTTCTCAGCGGTGATGGGGCTTCTTATGGCTTCCAGTACTCCCGCGCT 8226
 Qy 2641 GlnArgValGluPheLeuLeuLysAlaTrpAlaGluLysIleAspProMetGlyPheSer 2660
 Db 8227 CAACGGGTGGAGTTTCTTTGAAGGCGATGGCGGAAAGAACCCCTATGGGTGTTTCG 8286
 Qy 2661 TyrAspThrArgCysPheAspSerThrValThrGluArgAspIleArgThrGluGluSer 2680
 Db 8287 TATGATACCCGATGCTTCGACTCAACCGTCACTGAGAGAGATATCAGACTGAGGAATCC 8346
 Qy 2681 IleTyrArgAlaCysSerLeuProGluGluAlaHisThrAlaIleHisSerLeuThrGlu 2700
 Db 8347 ATATATCTGGCTGCTCCCTCGCGAGGCGCGCACTGCCATACATTCGCTGACTGAG 8406
 Qy 2701 ArgLeuTyrValGlyProMetPheAsnSerIleGlyGlnThrCysGlyTyrArgArg 2720
 Db 8407 AGACTTTTACGTGGAGGCGCCATGTTCAACAGCAAGGCGCGAGCTTCGGGTACAGCGGT 8466
 Qy 2721 CysArgAlaSerGlyValLeuThrThrSerMetGlyAsnThrIleThrCysTyrValLys 2740
 Db 8467 TGTCCGCCAGCGGGTCTCACTACTAGCATGGTAAACACCATACATGCTATGTAAAA 8526
 Qy 2741 AlaLeuAlaCysLysAlaAlaGlyIleIleAlaProThrMetLeuValCysGlyAsp 2760
 Db 8527 GCCCTAGCGCTTGCAGAGCGCGGGGATAGTTGCACCCCAATGCTGTGTGGCGGAC 8586
 Qy 2761 AspleuValIleSerGluSerGlnGlyThrGluGluAspGluArgAsnLeuArgAla 2780
 Db 8587 GACTTGGTTGTCACTCAGAAAGCCAGGGGACTGAGGAGGAGCGGAAACCTGAGAGGCC 8646
 Qy 2781 PheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProArgProGluTyr 2800
 Db 8647 TTCAGGAGGCTATGACAGGATATTCGCCCTCTGTTGACCCCGCCAGCCGGAATAT 8706
 Qy 2801 AspleuGluLeuIleThrSerCysSerSerAsnValSerValAlaLeuGlyProGlnGly 2820
 Db 8707 GACCTGGAGCTGATAACATCTTGTCTCCTCAATGTGTCTGTGGCACTCGGCCCCACAGGC 8766
 Qy 2821 ArgArgArgTyrTyrLeuThrArgAspProThrThrProIleAlaArgAlaAlaTrpGlu 2840
 Db 8767 CGCCGAGACTACTACCTGACAGAGACCTTACCACTCCAAATCGCCCGGGCTGCTGGGAA 8826
 Qy 2841 ThrValArgHisSerProValAsnSerTrpLeuGlyAsnIleIleGlnTyrAlaProThr 2860
 Db 8827 ACAGTTAGACACTCCCTCTGCTCACTATGGCTGGGAAACATCATCCAGTACGCTCCAACC 8886
 Qy 2861 IleTrpAlaArgMetValLeuMetThrHisPhePheSerIleLeuMetAlaGlnAspThr 2880
 Db 8887 ATATGGGTTCGAATGGTCTGATGACACATTTCTTCTCCATTTCTCATGGGCCAAGACACT 8946
 Qy 2881 LeuAspGlnAsnLeuAsnPheGluMetTyrGlyAlaValTyrSerValSerProLeuAsp 2900
 Db 8947 CTGGACAGAACCTTAACTTTGAATGTACGGATTCGGGTGATTCGGTGGAGCCCTTTGGAC 9006
 Qy 2901 LeuProAlaIleIleGluArgLeuHisGlyLeuAspAlaPheSerLeuHisThrTyrThr 2920
 Db 9007 CTCCAGCCATAATTTGAAGGCTACACGGGCTTACGCGCTTCTCTCTGCACACATACACT 9066
 Qy 2921 ProHisGluLeuThrArgValAlaSerAlaLeuArgLysLeuGlyAlaProProLeuArg 2940
 Db 9067 CCCCACGAACTAACCGCGGTGGCTTCAGCCCTCAGAAAACTTGGGGGCGCCACCCCTCAGG 9126
 Qy 2941 AlaTrpLysSerArgAlaArgAlaValArgAlaSerLeuIleSerArgGlyGlyArgAla 2960
 Db 9127 ACGTGGAGAGCTCGGCGCGGTGAGTTAGGCGCTCCCTCATCTCCATGGGGGGGAGAGCG 9186
 Qy 2961 AlaValCysGlyArgTyrLeuPheAsnTrpAlaValLysThrLysLeuLysLeuThrPro 2980
 Db 9187 GCCATTTCCGGTCCGGTATCTCTTCAACTGGGCGGTGAAAAACCAAGCTCAAACTCACTCCA 9246
 Qy 2981 LeuProGluAlaArgLeuLeuAspLeuSerTrpPheThrValGlyAlaGlyGly 3000
 Db 9247 TTGCGGAGGACGCCCTCTCGATCTATCCAGTTGTTTACCGTCCGCGCGCGGGGGC 9306

Qy 3001 AspIleTyrHisSerValSerArgAlaArgProArgLeuLeuPheGlyLeuLeuLeu 3020
 Db 9307 GACATTTATCACAGCGTGTGCGTGCCCGACCCCGCTTGTGTACTCTTTAGCCTACTCTTA 9366
 Qy 3021 LeuPheValGlyValGlyLeuPheLeuLeuProAlaArg 3033
 Db 9367 CTTTCCGTAGGGGTAGGCCTTCTCTACTTCCCGCTCGA 9405

Search completed: February 12, 2004, 06:41:12
 Job time : 24980 secs

QY 1901 IleCysAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrpMet 1920
 DB 6007 ATCTGGCGGCCATCTCTGGCGGCCACCTGGGGCCGGGAGCGCGGTTCAGTGGATG 6066
 QY 1921 AsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaProThrHisTyrValThr 1940
 DB 6067 AACAGGCTTATGCTTTGCTTCAGAGGAACACCGTGGCCGCCACCTACCTACGTCAGC 6126
 QY 1941 GluSerAspAlaSerGlnArgValThrGlnLeuGlySerLeuThrIleThrSerLeu 1960
 DB 6127 GAGTCGGATGCGTCGACGCTGTGACCAACTGCTTGGCTCCCTTACTATTAAGTACCTA 6186
 QY 1961 LeuArgArgLeuHisAsnTrpIleThrGluAspCysProIleProCysGlyGlySerTrp 1980
 DB 6187 CTCAGAAGACTCCACAACTGGATTACTAGAGACTGCGCTATCCATCCATGGCGCGGTTCGTG 6246
 QY 1981 LeuArgAspValTrpAspTrpValCysThrIleLeuThrAspPheLysAsnTrpLeuThr 2000
 DB 6247 CTCGGGACGTGTGGGACTGGGTTGACCATCTTAACAGACTTCAAAACTGGCTGACC 6306
 QY 2001 SerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTyrLys 2020
 DB 6307 TCCAGTTATTCCAAAGATGCCGGCTCCCTTTATCTTTGCCAAAGGGGTACAGG 6366
 QY 2021 GlyValTrpAlaGlyThrGlyIleMetThrThrArgCysProCysGlyAlaAsnIleSer 2040
 DB 6367 GGTGTCTGGGCTGGCACTGTATCATGACACACAGATGCTTTGGCGGCGCAACATCTCT 6426
 QY 2041 GlyAsnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTrp 2060
 DB 6427 GGTAAATGTCCTGGCTGGCTCTATGAGAAATACAGGGCCCTAAGACCTGCGATGAACACTTG 6486
 QY 2061 GlnGlyThrPheProIleAsnCysTyrThrGluGlyGlnCysValProLysProAlaPro 2080
 DB 6487 CAGGGAGACCTTCTCTATCAATGTACAGGGGGCAGTGGCTGGCGAGCGCGGCCCA 6546
 QY 2081 AsnPheLysValAlaIleTrpArgValAlaAlaSerGluTyrAlaGluValThrGlnHis 2100
 DB 6547 AACTTCAAGACCGCCATCTGGAGGGTGGGGCTCAGAGTACGCGAGGTGACGACAGCAC 6606
 QY 2101 GlySerTyrHisTyrIleThrGlyLeuThrThrAspAsnLysValProCysGlnLeu 2120
 DB 6607 GGTGTACTCTCTATATAACAGAAITGACCACTGACAATGTGAAGGTTCCTGCGCAACTA 6666
 QY 2121 ProSerProGluPhePheSerTrpValAspGlyValGlnIleHisArgPheAlaProThr 2140
 DB 6667 CTTTCTCCAGATTTTCTCTGGGTGGAGCGGGTGCAATCCATAGGTTCCTCCACG 6726
 QY 2141 ProLysProPhePheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal 2160
 DB 6727 CCAAAGCGTTTTTCCGGGATGAGGTCTGTTCCAGCGTTGGGCTCAATTCATTTGTCGTC 6786
 QY 2161 GlySerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThr 2180
 DB 6787 GGTCTCAGCTTCTTGTGATCTGAGCGCTGACACTGACGTATTGATGTCCTGCTAAACA 6846
 QY 2181 AspProSerHisIleThrAlaGluThrAlaAlaArgArgLeuAlaArgGlySerProPro 2200
 DB 6847 GATCCATCCCATATTACGGCGGAGCGCGAGCGCGGCTTACGCGGGGGGTACACCCCG 6906
 QY 2201 SerGluAlaSerSerAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThr 2220
 DB 6907 TCCGAGGCAAGCTCTCCACGACCGACCTATCGGCACCATCGTTGGTGGCTCCACCTGCACC 6966
 QY 2221 ThrHisGlyLysAlaTyrAspValAspMetValAspAlaAsnLeuPheMetGlyGlyAsp 2240
 DB 6967 ACCCAGCGCAGACCCCATGATGTGGACATGGTGGATGCAACCTGTTTCTTGGGGGGCAT 7026
 QY 2241 ValThrArgIleGluSerGlySerLysValValValLeuAspSerLeuAspProMetVal 2260
 DB 7027 GTGACCGGATAGTCCGAGTCCAAAGTGGTTGTTCTGGACTCTCTCGACCAATGGCC 7086
 QY 2261 GluGluArgSerAspLeuGluProSerIleProSerGluTyrMetLeuProLysLysArg 2280

DB 7087 GAAGAAAGGAGCGATCACGAGCCCTCGATACCATCAGACTACTTGTCTCCCAAGACGAG 7146
 QY 2281 PheProProAlaLeuProAlaTrpAlaArgProAspTyrAsnProProLeuValGluSer 2300
 DB 7147 TTCGGCCAGCTTACCGGCTGGGACGCGCTTGAGTCAACCCACCGCTTGTGGAAACCG 7206
 QY 2301 TrpLysArgProAspTyrGlnProAlaThrValAlaGlyCysAlaLeuProProProArg 2320
 DB 7207 TGGAAAGGGCAGATTACCAACGCGCGCGGTTCGGGCTGTGCTCTCCCGCCCTTAAG 7266
 QY 2321 LysThrProThrProProProArgArgArgThrValGlyLeuSerGluAspSerIle 2340
 DB 7267 AAAACCCCGACCCCGCCCAAGGAAACCCCGACAGTGGTCTGAACGAGACACTATA 7326
 QY 2341 GlyAspAlaLeuGlnGlnLeuAlaIleLysSerPheGlyGlnProProProSerGlyAsp 2360
 DB 7327 GCAGATGCCCTCCAGCAGCTGGCCATTAAACCTTTGGCCAGCCCTCCAAAGCGGGAT 7386
 QY 2361 SerGlyLeuSerThrGlyAlaGlyAlaAspSerGlySerGlnThrProProAspGlu 2380
 DB 7387 TCAGGCTCTCTACGGGGGCGACGCGCGCGATTCCGGCGGTTCAGACGCGCTCTGATGAG 7446
 QY 2381 LeuAlaLeuSerGluThrGlySerIleSerSerMetProProLeuGluGlyGluLeuGly 2400
 DB 7447 TTGGCTCATTTGGAGACAGGTTCCATCTCTCATGCCCGCTCGAGGGGAGCGCTGG 7506
 QY 2401 AspProAspLeuGluProGluGlnValGluProGlnProProProGlnGlyValAla 2420
 DB 7507 GACCCAGACCTCGAGCTCTGGGCGAGTAGAGCTTGAATCTCCCGCCCGAGGGGGAGGTA 7566
 QY 2421 AlaProGlySerAspSerGlySerTrpSerThrCysSerGluGluAspAspSerValVal 2440
 DB 7567 GCTCCGGCTCGACTCGGGGCTCTGCTTCTACTTCTCGAGAGGAGACGCTCCGCTCGT 7626
 QY 2441 CysCysSerMetSerTyrSerTrpThrGlyAlaLeuIleThrProCysSerProGluGlu 2460
 DB 7627 TGCTGCTCATCTCATCTCTTGGACCGGGCTCTTAATACTCTTGTAGCCCGGAGAG 7686
 QY 2461 GluLysLeuProIleAsnProLeuSerSerSerLeuLeuArgTyrHisAsnLysValTyr 2480
 DB 7687 GAAAGTTGGCAATTAACCCCTTGAGCAACTCGCTGTTGCGATACCCACAAAGGTGTAC 7746
 QY 2481 CysThrThrThrLysSerAlaSerLeuArgAlaLysLysValThrPheAspArgMetGln 2500
 DB 7747 TGTTACTACAGCAAGAGCGCTCATTTGAGCGCTTAAAGGTAACTTTTGTATAGATGCAA 7806
 QY 2501 ValLeuAspSerTyrTyrAspSerValLeuLysAspIleLysLeuAlaAlaSerLysVal 2520
 DB 7807 GTGCTCGAGCCCATATACCTCAGTCTTAAAGGACATCAAGCTGGCGGCTCCCAAGTC 7866
 QY 2521 ThrAlaArgLeuLeuThrMetGluAlaCysGlnLeuThrProProHisSerAlaArg 2540
 DB 7867 AGCGCAAGGCTCTCCTCAGCTTGGAGGAGCATGCCAATTAAGCCACCCCTTCGCAAGA 7926
 QY 2541 SerLysTyrGlyPheGlyValAlaLysGluValArgSerLeuSerGlyArgAlaValAsnHis 2560
 DB 7927 TCCAAAGTATGGTTTGGGGCTAAGGAGGTCCGAGCTTGTCCGGGAGGCGCTCAACCAC 7986
 QY 2561 IleLysSerValTrpLysAspLeuLeuGluAspSerGluThrProIleProThrThrIle 2580
 DB 7987 ATCAAGTCGTGTGAGAGGACCTCTCGGAAGACTCACAACACCAATTCCTACGACCATC 8046
 QY 2581 MetAlaLysAsnGluValPheCysValAspProThrLysGlyGlyLysLysAlaArg 2600
 DB 8047 ATGGCAAAAAATGAGGTGTTCTGCGTGACCCCGCCACCAAGGGGGTAAAGAACGAGCTCG 8106
 QY 2601 LeuIleValTyrProAspLeuGlyValArgValCysGluLysMetAlaLeuTyrAspIle 2620
 DB 8107 CTTATCTGTTTACCTCGACTCGGTGTAGGGGTCTCGAGAAATAGCCCTTTTATGATGTC 8166
 QY 2621 ThrGlnLysLeuProGlnAlaValMetGlyAlaSerTyrGlyPheGlnTyrSerProAla 2640

Copied from 09540843 on 05/19/2004

Copied from 09540843 on 05/19/2004

BASE COUNT	1952 a	2790 c	2652 g	2021 t	1 others
ORIGIN					

	Alignment Scores:	
Pred. No.:	0	Length:
Score:	15347.00	Matches:
Percent Similarity:	97.00%	Conservative:
Best Local Similarity:	94.30%	Mismatches:
Query Match:	95.20%	Indels:
	14	Gaps:
DB:	15	0
		0
		9416

US-09-980-559-2 (1-3033) x AF238481 (1-9416)

1	MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgProGln	20
307	ATGAGCACAATCTCTAAACCTCAAGAGAAAAACCAAGAAACACATAACCGTCGCCACAA	366
21	AspValLysPheProGlyGlyGlyGlnIleValGlyValTyrlleuLeuProArgArg	40
:	:	:
367	AACGTGAAGTTTCGGCGCGCGCCACAGATCGTTGGCGGAGTACTTGTTCGCCGCACGG	426
41	GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly	60
427	GCCTCCAGTTGGGTGTGCGCGCGCAAGAAGAACTTCGGAGCGGTCCCGACCACTGGG	486
61	ArgArgGlnProIleProLysAspArgArgSerThrGlyLysSerTrpGlyLysProGly	80
487	AGGCGCGAGCTCATCTCTTAAGATCCGCGCTCCACATGCAAGTCTCTGGGGAACACAGA	546

81	TyrProTrrpProLeuTyrGlyAsnGluGlyLeuGlyTrrpAlaGlyTrrpLeuLeuSerPro	100	Qy
547	TACCCCTGGCCCTATACGGGAATGAGGAGCTCGCGTCGGCAGGATGGCTCTGTGTC	606	Db
101	ArgGlySerArgProSerTrrpGlyProAenAspProArgHisAetSerArgAsnValGly	120	Qy
607	CGAGGTCCTCCGTCCTCTTGGGGGCCCCAATGACCCCGGCATAGAGTCGGCAACGTGGGT	666	Db
121	LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProValVal	140	Qy
667	AAGGTCATCATACCCCTAAGCTGGCGCTTGGCCGACCTCATGGGGTACATCCCGCTCGTA	726	Db
141	GlyAlaProLeuGlyGlyValalabargalaleuAlaHisGlyValargValleuGluAsp	160	Qy
727	GGCGCCCGCTGGTGGCGTGGCCAGAGCTCTCGCGCACGGCGTGAGAGTCTCTGGAGGAC	786	Db
161	GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla	180	Qy
787	GGGGTTAATTTGCAACAGGGAACCTTACTGTGTGCTCTCTTCTTCTATTTTCTGTGGCC	846	Db
181	LeuLeuSerCysIleThrThrProValSerAlaAlaGluValIleAsnIleSerThrGly	200	Qy
847	CTACTGTCTTCATCACCATTCCGGTCTCTGTGTCCAAGTCAAGACACCAAGTAAACAGC	906	Db
201	TyrMetValThrAsnAspCysThrAsnAspSerIleThrTrrpGlnleuGlnalalaVal	220	Qy
907	TATATGGTGACTAATGATTTGCCAACACACAGACATTACCTGGCAGGTTCCAGGACGGT	966	Db
221	LeuHisValProGlyCysValProCysGluIysValGlyAsnAlaSerClnCystrIle	240	Qy
967	CTCCACGTCCTCCGGGTGGCTCCGTCGCGAAGTGGAAATACATCATCGGTCTGGATA	1026	Db
241	ProValSerProAsnValAlaValGlnArgProGlyAlaLeuThrGlnGlyLeuArgThr	260	Qy
1027	CCGGTCTCACCAACAGTGTGTGGCAGCGCGCGCCCTCACGCAAACTTGGCGAGC	1086	Db
261	HisIleAspMetValValMetSerAlaThrLeuCysSerAlaLeuTyrValGlyAspLeu	280	Qy
1087	CACATCATATGTTGTGATGTCCGCCACGCTGTGTCGCACTCTACGTGGGGATCTC	1146	Db
281	CysGlyGlyValMetLeuAlaAlaGlnMetPheIleValSerProGlnHisIleTrrpPhe	300	Qy
1147	TGCGGGGGGTGATGTCTGCACGCCAAATGTTCAATGTCTCGCGGAGCGCCACTGGT	1206	Db
301	ValGlnAspCysAsnCysSerIleTyrProGlyThrIleThrGlyHisArgMetAlaTrrp	320	Qy
1207	GTGCAAGATGCAATGTCTCCATCTACCTTGGCACCATCTATGGACACCATGTATGGATGG	1266	Db
321	AspMetMetMetAsnTrrpSerProThrAlaThrMetIleLeuAlaTyrAlaMetArgVal	340	Qy
1267	GATATGATGATGAATGGTTCGCCCCACGACCAACCATGATCTTGGCGTACGCGATGCGCAT	1326	Db
341	ProGluValIleIleAspIleIleSerGlyAlaHisTrrpGlyValMetPheGlyLeuAla	360	Qy
1327	CCCGAGGTATCATGACATCATGACGCGGGGTCTACTGGGGCGTCAATGTTCCGCCCTAGCC	1386	Db
361	TyrPheSerMetGlnGlyAlaTrrpAlaLysValValIleLeuLeuLeuAlaAlaGly	380	Qy
1387	TACTTCTCCATGAGGAGCGTGGGGGAAGTGTGTTGTATCTCTTGTGTGACCGCTGGG	1446	Db
381	ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr	400	Qy
1447	GTGGACGCGGAACCCAAACAGTTGGGGGTGCAACCGGGCGAGTGTCTTGGGCTTCACC	1506	Db
401	SerLeuPheAspMetGlyProArgGlnIlyIleGlnLeuValAsnThrAsnGlySerTrrp	420	Qy
1507	AGCCTCTTCAACGGTGGCCCTCAACAGCAGATTCAGTCTCATCAACCAATATGCAATGG	1566	Db
421	HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer	440	Qy
1567	CACATCAACCGCACTGCCCTGAATTCGAATGACTCTTGAACACCGGCTTATCGCGTCC	1626	Db

Db	7901	AGCCCAAGGCTCCCTCACTTTGGAGGCGGTGCAATTGATCCACCCCACTTCTGCAAGA	7960	Db	8981	CTGGACGACAGAACCTCAACTTTGAAATGTACGGATCGGTGACTCTCGGTAGTCTCTTTGGAC	9040
Qy	2541	SerLysTyrGlyPheGlyAlaLysGluValArgSerLeuSerGlyArgAlaValAsnHis	2560	Qy	2901	LeuProAlaIleleGluArgLeuHisGlyLeuAspAlaPheSerLeuHisThrThr	2920
Db	7961	TCCAAATATGGTTTGGGCAAGAGAGGTTCGACGTTGTCCGGAGGCGGTCAACCAC	8020	Db	9041	CTCCAGCCATATTCAGAGTTACACGGCTTGACGCTCTCTCTGACACATCTCT	9100
Qy	2561	IleLysSerValTrpLysAspLeuLeuGluAspSerGluThrProLeuThrIle	2580	Qy	2921	ProHisGluLeuThrArgValAlaSerAlaLeuArgLysLeuGlyAlaProLeuArg	2940
Db	8021	ATCAAGTCCGTGTGAAGGACCTCTCGAAGACTCACAAACACCAATTCCTACGACCATC	8080	Db	9101	CCCCAGCACTGACGCGGTGGCTTCGGCCCTCAGAAACCTTGGGGCGCCACCACTCAGA	9160
Qy	2581	MetAlaLysAsnGluValPheCysValAspProThrLysGlyLysAlaAlaArg	2600	Qy	2941	AlaTrpLysSerArgAlaArgAlaValArgAlaSerIleSerArgGlyArgAla	2960
Db	8081	ATGCCCAAAATAGGTTCTCTGGTGGACCCCACTAAGGGGGGCAAGAAAGCAGCTCGC	8140	Db	9161	GGCTGGAAGAGTCGGCGGTGAGTTAGGGCGTCTCATCTCCGAGGGGGAGACGC	9220
Qy	2601	LeuIleValTyrProAspLeuGlyValArgValCysGluLysMetAlaLeuTyrAspIle	2620	Qy	2961	AlaValCysGlyArgTyrLeuPheAsnTrpAlaValLysThrLysLeuLysLeuThrPro	2980
Db	8141	CTTATCTTTACCCCGACCTCGGCGTCAGGGTTCGCGAAGATGGCCCTTTATGATGTC	8200	Db	9221	GCATTTGTGGCGGTATCTTTCAACTGGCGGTGAAGACTAAGCTCAAACTCACTCCA	9280
Qy	2621	ThrGlnLysLeuProGlnAlaValMetGlyAlaSerTyrGlyPheGlnTyrSerProAla	2640	Qy	2981	LeuProGluAlaArgLeuLeuAspLeuSerSerTrpPheThrValGlyAlaGlyGly	3000
Db	8201	ACACAAAGCTTCTCAGCGGTGATGGGGCTTCTATGGCTTCCAGTACTCCCGCGCT	8260	Db	9281	TTGCGGAGGACGCGCTTCTGGATCTATCAGTTGTTTACCGTCGCGCGCGGGGCG	9340
Qy	2641	GlnArgValGluPheLeuLysAlaTrpAlaGluLysLysAspProMetGlyPheSer	2660	Qy	3001	AspIleTyrHisSerValSerArgAlaArgProArgLeuLeuLeuPheGlyLeuLeu	3020
Db	8261	CAACGGGTGGAGTTTCTTGAAGGCAATGGCGGATGAAGAGAACCATGGGGTTCTCG	8320	Db	9341	GACATTTATCACAGCGTGTGCGCGCCGACCCCGCTTATTACTCTTAGCTTACTCTA	9400
Qy	2661	TyrAspThrArgCysPheAspSerThrValThrGluArgAspIleArgThrGluGluSer	2680	Qy	3021	LeuPheValGlyValGlyLeuPheLeuLeuProAlaArg	3033
Db	8321	TATGATACCGAGTCTTTGACTCACTCACTGTCACGAGAGACATTAAGACTGAGGAGTCG	8380	Db	9401	CTTTCTGATGGGTAGGCTCTCTTACTCTCCCGCTCGG	9439
Qy	2681	IleTyrArgAlaCysSerLeuProGluGluAlaHisThrAlaIleHisSerLeuThrGlu	2700	RESULT 15			
Db	8381	ATATACAGGCTGCTCTCTCCGCGAGGAGGCTCACACTGCCATACACTCGCTGACTGAG	8440	LOCUS	AF238481	9416 bp	RNA linear VRL 26-MAR-2000
Qy	2701	ArgLeuTyrValGlyGlyProMetPheAsnSerLysGlyGlnThrCysGlyTyrArgArg	2720	DEFINITION	Hepatitis C virus 2a polyprotein gene, complete cds.		
Db	8441	AGACTCTACGTGGGAGGCGCATGCTCAACAGCAAGGCGCGCTCGGGGTACAGGCGT	8500	ACCESSION	AF238481		
Qy	2721	CysArgAlaSerGlyValLeuThrThrSerMetGlyAsnThrIleThrCysTyrValLys	2740	VERSION	AF238481.1	GI:7329200	
Db	8501	TGCGCGCCAGCGGGGTCTTACACATGAGTGGGAATACCATCATCTATGTGAAA	8560	KEYWORDS			
Qy	2741	AlaLeuAlaAlaCysLysAlaAlaGlyIleleAlaProThrMetLeuValCysGlyAsp	2760	SOURCE	Hepatitis C virus		
Db	8561	GCCTAGCGCTTGCAGAGCTGCGGGATAGTGCACCTACATGTGTATCGGCGAT	8620	ORGANISM	Hepatitis C virus		
Qy	2761	AspLeuValLysSerGluSerGlnGlyThrGluGluAspGluArgAsnLeuArgAla	2780	REFERENCE	1 (bases 1 to 9416)		
Db	8621	GACTTGGTTGTCATCTCAGAAAGCCAGGAGCTGAGGAGCAGCGGAACCTGAGAGCT	8680	AUTHORS	Itakura,J., Nagayama,K., Enomoto,N., Kurosaki,M., Watanabe,H. and Sato,C.		
Qy	2781	PheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProArgProGluTyr	2800	TITLE	Full length cDNA sequence of HCV genotype 2a		
Db	8681	TTCAAGGAGCTATGACAGGATATTCGCGCCCTCTGTCGATCCCGCAGACCGGAATAT	8740	JOURNAL	Unpublished		
Qy	2801	AspLeuLeuIleThrSerCysSerSerAsnValSerValAlaLeuGlyProGlnGly	2820	AUTHORS	Itakura,J., Nagayama,K., Enomoto,N., Kurosaki,M., Watanabe,H. and Sato,C.		
Db	8741	GACTTGGAGCTGATACATCTTGTCTCTCAAAATGTCTGTGACCTCGGCCACAGGCG	8800	TITLE	Direct Submission		
Qy	2821	ArgArgArgTyrTyrLeuThrArgAspProThrThrProIleAlaArgAlaAlaTrpGlu	2840	JOURNAL	Submitted (24-FEB-2000) Second Department of Internal Medicine, Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-Ku, Tokyo 113-8519, Japan		
Db	8801	CGCGCAGATACCTGACGACGAGATCTTACCACTCCATCGCCCGGGCTGCTGGAA	8860	FEATURES	Location/Qualifiers		
Qy	2841	ThrValArgHisSerProValAsnSerTrpLeuGlyAsnIleleGlnTyrAlaProThr	2860	source	1. .9416		
Db	8861	ACAGTCAGACACTCCCTCTGCAATTCATGCTCGGAAACATCATCTCAGTACGCCCAACC	8920		/organism="Hepatitis C virus"		
Qy	2861	IleTrpAlaArgMetValLeuMetThrHisPhePheSerIleLeuMetAlaGlnAspThr	2880		/mol_type="genomic RNA"		
Db	8921	ATATGGGTTGGATGGTCTTGATGACACATCTTCTTCCATTTTATGGCCCAAGACACC	8980		/db_xref="taxon:11103"		
Qy	2881	LeuAspGlnAsnLeuAsnPheGluMetTyrGlyAlaValTyrSerValSerProLeuAsp	2900		/note="genotype: 2a"		
					307. .9408		
					/codon_start=1		
					/product="polyprotein"		
					/protein_id="AAF5940.1"		
					/translation="MSTNPKPQRTKRNTRRPNVKKFGGQIVGGVYLLPRRGRL		
					SVRAKTSRSQPRGRQVITLTCGADLMGYIPVVGAPLGGVARALAGVRVLED		
					GYNFATNLPCCFSIFLLALLSCITIPVSAVOKNTSNVMTNDSCNLSITWQOD		
					AVLHPGCVPEKVENTSRCLWIPSPNVAVRORGLATQNLRLTHIDNMVNSATLCSALY		
					VSDCCGVMLAAQMFIVSPRHFWFOECNCSIPGTTTGRHAWDMNMWNSPTTMMIL		
					AYAMRIPEYIDIMSGAHGVMFLAYFSMQGAAKVVVILLTLAGVDAETQTVGAT		
					GRSLVGLFTSLFTRGPQQQLINTNGSHINRTALNCNDLSLNTGTFTASLFTYNNFNS		

Db 3521 GGGGAAATCCAGTCTTGTCCACAGTCAGTCAGTCTCTCCGGAACATCCATTTCCGGG 3580
 QY 1081 ValLeuThrValThrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro 1100
 Db 3581 GTCTTGGGACTGTTTACACAGGAATCGGCAATAAGACCTTAGCCGGCTCACGGGGCCG 3640
 QY 1101 ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTyrProSerProGly 1120
 Db 3641 GTCACACAGATGATTCAAGTCCGAGGGGACTTGGTAGGGTGGCCAGTCCTCCGGG 3700
 QY 1121 ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArgAsn 1140
 Db 3701 ACCAAATCTTTGGAGCGTGCAGCTGTGGAGCGGTGCACCTATACCTGTGTCCAGCAAC 3760
 QY 1141 AlaAspValLeuProAlaArgArgGlyAspLysArgGlyAlaLeuLeuSerProArg 1160
 Db 3761 GCTGATGTCTATCCCGCTCGAAGACCGGGGCAAGCGGGGAGCGTACTCTCCCGAGG 3820
 QY 1161 ProLeuSerThrLeuLysGlySerSerGlyGlyProValLeuCysProArgGlyHisAla 1180
 Db 3821 CCTCTTTCGACCTGAAGGATCATCGGGGGACCGGTGCTCTGCCCTAGGGGCCAGCT 3880
 QY 1181 ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerLeuAspPheile 1200
 Db 3881 GTCGGGATCTCCGGGCGAGCTGTGCTCTCGGGGGTGGCTAAGTCCATAGATTTTCATC 3940
 QY 1201 ProValGluThrLeuAspLeuValThrArgSerProThrPheSerAspAsnSerThrPro 1220
 Db 3941 CCGCTTGAGCGCTGCACATCATGTCAGCGGTCTCCACCTTTTAGCCACCAACACACCA 4000
 QY 1221 ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLys 1240
 Db 4001 CCAGCTGTGCCCCAGACCTATCAGTTCGGGTACTTGCATGCCGCCCAACTGGCAGTGGAAAG 4060
 QY 1241 SerThrLysValProValAlaTyrAlaGlnGlyTyrLysValLeuValLeuAsnPro 1260
 Db 4061 AGCACCAAGTCCCTGTGCGATATGCCGCCAGGGGTACAAAGTATTAGTGTCTTAACCCC 4120
 QY 1261 SerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerLysAlaHisGlyLeuAsnPro 1280
 Db 4121 TCGGTGGCTGCCACTCTGGGGTTTGGGGGTACTCTGTCGAAGGCACATGCGATCAATCCC 4180
 QY 1281 AsnIleArgThrGlyValArgThrValThrThrGlyAlaProIleThrTyrSerThrTyr 1300
 Db 4181 AACATCAGACTCGAGTCAAGTGTGACGACCGGGGGGCCCATCATCAGTACTCCAGTAT 4240
 QY 1301 GlyLysPheLeuAlaAspGlyGlyCysAlaGlyAlaTyrAspLeuLeuLysCysAsp 1320
 Db 4241 GGTAAATTCCTGCTGATGGGGCTGCTCAGGGCGGCGCTATGACATCATATCGCAT 4300
 QY 1321 GluCysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGlnAla 1340
 Db 4301 GAATGCCAGCTGTGACCGCTACCACTCTCGGCATCGGAACGGTCTTTCGACCAAGCT 4360
 QY 1341 GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProProGlySerValThr 1360
 Db 4361 GAGACAGCCGGGTCAAGCTCACTGACTGCGCCACCGGCCACCGCCCCCGGGTGGTGACA 4420
 QY 1361 ThrProHisProAsnIleGluGluValAlaLeuGlyGlnGluGlyLeuIleProPheTyr 1380
 Db 4421 ACTCCCATCCCAACATAGAGAGATAGCCCTCGGGCAGGGGTGAGATCCCTCTCTAT 4480
 QY 1381 GlyArgAlaIleProLeuSerTyrIleLysGlyIleArgHisLeuIlePheCysHisSer 1400
 Db 4481 GGGAGGGCGATTCCCTGTCTCATCATCAAGGGGAGGAGACACTTAATCTTCTGCCACTCA 4540
 QY 1401 LysLysCysAspGluLeuAlaAlaAlaLeuArgGlyMetGlyLeuAsnSerValAla 1420
 Db 4541 AAAAAAAGTGTGACGAGCTCGCGGGCCCTTCGGGGCATGGGCTTGAACCGCTGTGGCA 4600
 QY 1421 TyrTyrArgGlyLeuAspValSerValIleProThrGlnGlyAspValValValAla 1440

Db 4601 TACTACAGAGGGTGGACGCTCTCCATAATACCAACTCAAGAGGATGTAGTGGTCTGCGCC 4660
 QY 1441 ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnVal 1460
 Db 4661 ACCGAGCTCTCATGACGGGGTTCGTGGAGACTTTGACTCTGTGTATCAGCTCAACGTA 4720
 QY 1461 AlaValThrGlnValValAspPheSerLeuAspProThrPheThrIleThrGlnIle 1480
 Db 4721 GGGGTCACTCAAGTCTGGACTTACGCTGGAGCCACCTTCACTATTAACACAGACT 4780
 QY 1481 ValProGlnAspAlaValSerArgSerGlnArgGlyArgThrGlyArgGlyArgLeu 1500
 Db 4781 GTCCTCTCAAGACGCTCTCTCAGTAGCAGCGCGGGGGCTACAGGTAGAGAGACTG 4840
 QY 1501 GlyIleTyrArgTyrValSerThrGlyGluArgAlaSerGlyMetPheAspSerValVal 1520
 Db 4841 GGCACCTTACAGGTACGTTTCCACTGGAGAGCGAGCTCAGGAATGTTTACAGTGTAGTG 4900
 QY 1521 LeuCysGluCysTyrAspAlaGlyAlaAlaTyrTyrGluLeuThrProSerGluThrThr 1540
 Db 4901 CTCTGTGAGTGTACGACCGAGGGCTGCATGGTATGAGCTTACGCCAGTGGAGACTACC 4960
 QY 1541 ValArgLeuArgAlaTyrPheAsnThrProGlyLeuProValCysGlnAspHisLeuGlu 1560
 Db 4961 GTCAGCTTTGGAGCGTATTTCATACGCTGGCTTGCCTGTGTGCCAAGACCACTTGAG 5020
 QY 1561 PheTyrGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThr 1580
 Db 5021 TTTTGGAGGCGAGTTTTCACCGGCTTACACACATAGATGCCACCTCTCTTCCCAGACA 5080
 QY 1581 LysGlnSerGlyGluAsnPheAlaTyrLeuThrAlaTyrGlnAlaThrValCysAlaArg 1600
 Db 5081 AAACAATCGGGGAGAAATTCGGTACTTAGTAGCTATCAGCCACAGCTGTGGCTAGA 5140
 QY 1601 AlaLysAlaProProProSerTrpAspValMetTrpLysCysLeuThrArgLeuLysPro 1620
 Db 5141 GCCAAAGCCCCCCCCGCTTGGGAGCGTCATGTGNAATGCTTGACTCGACTCAAGCCC 5200
 QY 1621 ThrLeuValGlyProThrProLeuThrArgLeuGlySerValThrAsnGluValThr 1640
 Db 5201 ACAGCTGTGGGGCTTACACCTCTCTGTACCGCTTGGGCTCTGTATCAATAGGTGACC 5260
 QY 1641 LeuThrHisProValThrLysTyrIleAlaThrCysMetGlnAlaAspLeuGluValMet 1660
 Db 5261 CTTCACACATCTGTGACAAAATACATCGCCACATGCATGCAAGCTGACCTTGGGTCTATG 5320
 QY 1661 ThrSerThrTrpValLeuAlaGlyGlyValLeuAlaAlaValAlaAlaTyrCysLeuAla 1680
 Db 5321 ACCAGCAGCTGGGTCTGTGGCTGGGGAGTCTTGGCAGCGCTCGCCGCTATTGGTGGCG 5380
 QY 1681 ThrGlyCysValCysIleGlyArgLeuHisIleAsnGlnArgAlaValAlaPro 1700
 Db 5381 ACCGGGTGTGTTCATCATCATCGCGGTGGCATCAACAGCGAGCTGTGTGCTGCTCGG 5440
 QY 1701 AspLysGluValLeuTyrGluAlaPheAspGluMetGluGluCysAlaSerArgAlaAla 1720
 Db 5441 GACACAGAGGTCTCTCTATGAGGCTTTTGNATGAGATGGGAAGATGTGCTCTAGAGCGGT 5500
 QY 1721 LeuIleGluGlyGlnArgIleAlaGluMetLeuLysSerLysIleGlnGlyLeuLeu 1740
 Db 5501 CTCTATTGGAGGGGCGAGCGGATGGCCGAGATGTTGNAAGTCTTAAGATCCCAAGGCTTATTG 5560
 QY 1741 GlnGlnAlaSerLysGlnAlaGlnAspIleGlnProThrValGlnAlaSerTrpProLys 1760
 Db 5561 CAGCAAGGCTCTAAACAGGCCAGGACATACACCTGTGTGTGAGGCTCTGTGGGCCAAG 5620
 QY 1761 ValGluGlnPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeuAla 1780
 Db 5621 GTGGAGCAGTCTTGGGCCAAACACATGTGAATTTTATTAGCGGCTTCAATACCTCGCA 5680
 QY 1781 GlyLeuSerThrLeuProGlyAsnProAlaValAlaSerMetMetAlaPheSerAlaAla 1800
 Db 5681 GGACTGTCAACACTGCCAGGGAACCTGTGTGGCTTCCATGATGGCATTCAGTGGCGCT 5740

Qy	341	ProGluValIleIleAspIleIleSerGlyAlaHisIleIleGlyValMetPheGlyLeuAla	360	Qy	701	ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrLysTyrIleValArgTrrp	720
Db	1361	CCGAGGTATCATAGACATATATAGCGGGCTCATCTGGGGTGTATGTTTCGGCCCTGGCC	1420	Db	2441	GTGGAGGTACATACATGTATGGCTGTACCTGGCCCTTACAAATATATGCTTCGATGG	2500
Qy	361	TyrPheSerMetGlnGlyAlaIleAlaLysValValIleLeuLeuAlaGly	380	Qy	721	GluTrpValIleLeuLeuPheLeuLeuAlaAspAlaArgValCysAlaCysLeuTrrp	740
Db	1421	TACTTCTATGAGGGAGGTGGGCAAGGTGTGTGTATCTTCTGTCTGGCTGGG	1480	Db	2501	GAGTGGGTGTACTTCTATTTCTGTCTAGCGGAGCCAGGGTCTGCGCCCTGTGTGTGG	2560
Qy	381	ValAspAlaArgThrHisThrValGlyCysAlaAlaGlnThrThrGlyArgLeuThr	400	Qy	741	MetLeuLeuLeuGlyGlnAlaGluAlaAlaLeuGluLysLeuValIleLeuHisAla	760
Db	1481	GTGGACGGGGCACCCATACAACTGGGGCTGTTCGGCGAGCAACGCTAGAGCCCTCGCC	1540	Db	2561	ATGCTCATCTTGTGGCCAGCTGAGCAGCAGCTAGAGAGCTGTGTCTTGCACGCT	2620
Qy	401	SerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTrrp	420	Qy	761	AlaSerAlaAlaSerCysAsnGlyPheLeuTyrPheValIlePhePheValAlaIleTrrp	780
Db	1541	AGTTTGTCTACTTCTGGACCTAGCAGCAATTCAGCTCAATTAACCAATGCGAGTTGG	1600	Db	2621	GCGAGCGCAGCTAGTTGCAATAGCTTCTGTATTTTGTCTATCTTTTCTGTGGCTGTGG	2680
Qy	421	HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer	440	Qy	781	TyrIleLysGlyArgValValProLeuAlaThrTyrSerLeuThrClyLeuTrrpSerPhe	800
Db	1601	CACATTAACCGCACTGCGCTGAAATTGCAATGACTCCCTGCAAACTGGCTTCACTCGGCC	1660	Db	2681	TACATCAAGGGTGGGCGAGTCCCTGGCTGTCTATTCCTCACTGGCTATGGCCCTTC	2740
Qy	441	LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg	460	Qy	801	SerLeuLeuLeuAlaLeuProGlnGlnAlaTyrAlaTyrAspAlaSerValHisGly	820
Db	1661	CTGATCTACAGGAACCGCTTCAACTGCTCAGGATGTCCCGAACGCTGTCCGTGTGCCGC	1720	Db	2741	TGCCTACTGCTTACGACTGCGCCCAACAGGCTTACGCTATGACGCACTGTGTGCACGA	2800
Qy	461	SerIleGluAlaPheArgValGlyTrrpGlyAlaLeuGlnTyrGluAspAsnValThrAsn	480	Qy	821	GlnIleGlyAlaAlaLeuLeuValMetIleThrLeuPheThrLeuThrProGlyTrrpLys	840
Db	1721	AATATCAGGCTTCCCGAGATAGGATGGGGCGCTTGCATACAGAGGATATGTCATTAAT	1780	Db	2801	CAGATAGGCGGCTTTTGTGTACTGATTACCTCTTTTACACTTACCCCGGGGTATAA	2860
Qy	481	ProGluAspMetArgProTyrCysTrrpHisTyrProProArgGlnCysGlyValSer	500	Qy	841	ThrLeuLeuSerArgPheLeuTrrpLeuCysTyrLeuLeuThrLeuGlyGluAlaMet	860
Db	1781	CCAGAGATATGAGACCATATGCTGGCACTATCCACCAACCACTGTGGCAATTGTCCCG	1840	Db	2861	ACCTCTCTCAGCGGAGTCTGTGGTGGCTGTCTCTCTGACCTTGGGGGAGGCGCATG	2920
Qy	501	AlaLysThrValCysGlyProValTyrCysPheThrProSerProValValGlyThr	520	Qy	861	ValGlnGluTrrpAlaProProMetGlnValArgGlyGlyArgAspGlyIleIleTrrpAla	880
Db	1841	GCGCAGTCAAGTGTGGCCCGCTACTTCTACCCCGCCAGTGGTAGTGGGCACA	1900	Db	2921	GTTTCAGAGTGGGCACACCCATGACGGCGGTGTGGCGGTGATGCATCATATGGGCC	2980
Qy	521	ThrAspArgLeuGlyAlaProThrTyrThrTrrpGlyGluAsnGluThrAspValPheLeu	540	Qy	881	ValAlaIlePheTrrpProGlyValValPheAspIleThrLysTrrpLeuLeuAlaValLeu	900
Db	1901	ACCAGACAGCGGAGGCGCTTCTTACAGTGGGAGAGATGATACAAATGCTTCTCTA	1960	Db	2981	GCCACCATCTCTGCGCGGCGTAGTGTGTATATAACCAAGTGGCTGTGTGGCGGTCTT	3040
Qy	541	LeuAsnSerThrArgProProLeuGlySerTrrpPheGlyCysTrrpMetAsnSerSer	560	Qy	901	GlyProAlaTyrIleLeuLysGlyAlaLeuThrArgValProTrrpPheValArgAlaHis	920
Db	1961	TTGAACAGCACTCGACCAACCGCTGGGTGATGTTGGATGCAGCTGGATGAATCCCACT	2020	Db	3041	GGCGCGCTTACCTCTGAGAGATGCTTTGACGAGCGTGGCTTCTGTCAGAGCTCAC	3100
Qy	561	GlyTyrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer	580	Qy	921	AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyGlyArgTrrpValGlnMet	940
Db	2021	GGTTTCAACCAAACTTGTGGCGCACCACTTGTCCGCAATTAGAGCTGACTTCAATACCAGC	2080	Db	3101	GCTCTGCTGAGATGTGCATATGGTGAAGCACTCGCGGGGGTAGTACGTCCAGATG	3160
Qy	581	ThrAspLeuLeuCysProThrAspCysPheArgLysHisProAspThrThrTrrpLys	600	Qy	941	AlaLeuLeuAlaLeuGlyArgTrrpThrGlyTrrpIleTrrpAspHisLeuThrProMet	960
Db	2081	ACGGACCTTATGCCCCCAGGATGTTTATAGGAAGATCTCTGATGCCACTTACATCAAA	2140	Db	3161	GCGCTATTAGCCCTTGGTAGATGGTCCGGCACTTACATCTATACCACTTACCCCTTATG	3220
Qy	601	CysGlySerGlyProTrrpLeuThrProArgCysLeuIleAspTrrpProTrrpArgLeuTrrp	620	Qy	961	SerAspTrrpAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIleIlePhe	980
Db	2141	TGTGTTCTGGCCCTTGGCTCAGCCAAAGTGCCTGGTGCAGTACCCCTACAGGCTCTGG	2200	Db	3221	TCGGATTGGGCTCTAGCGGCTTGGCGGACTTGGCGGTCTGTGGAGCTTATCATCTTC	3280
Qy	621	HisTyrProCysThrValAsnTyrThrIlePheLysIleArgMetTrrpValGlyVal	640	Qy	981	SerProMetGluLysLysValIleValTrrpGlyValaGluThrAlaAlaCysGlyAspIle	1000
Db	2201	CATTACCCCTCCACAGTTAACTATACCACTTCAAGGTAAAGGATGTACGTGGGGGGGT	2260	Db	3281	AGTCCGATGGAGAAAGTCACTGCTGGGAGCAGACGCTGCATCGCGGGGACATC	3340
Qy	641	GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp	660	Qy	1001	LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp	1020
Db	2261	GAGCACAGGCTCAGGGCTGCGTCAACTTCACTCGCGGGATCGTGTCACTTGGAGGAC	2320	Db	3341	TTATACGGACTTCTGTGTCCGCCGACTCGGTGGGAGATCTCTCTTGGCCCACTGAT	3400
Qy	661	ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGluTrrpAlaIleLeuPro	680	Qy	1021	GlyTrrpThrSerLysGlyTrrpSerLeuLeuAlaProIleThrAlaTrrpAlaGlnThr	1040
Db	2321	AGGACAGAGTCAACAGTCTCTTGTGTGCACTCCACCGAATGGGCGCTTGTGCC	2380	Db	3401	GGCTTACACCTCCAAAGGGGTGGAGCTTCTCGCCCCCATCTCTCTACGCCACGACACA	3460
Qy	681	CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisLeuHisGlnAlaIle	700	Qy	1041	ArgGlyLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGluGlnAla	1060
Db	2381	TGCTCTTCTCAGATCTGCGCGCTGTGTCGATGGCTTCTTCTACCTCCACCAAAATATC	2440	Db	3461	CGAGGTCTCTGGCGCCCATAGTAGCATGACGGGGGAGACAGACAGACAGCAGGCC	3520
Qy				Qy	1061	GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly	1080

FEATURES

Location/Qualifiers
 1. .9691
 /organism="Hepatitis C virus"
 /mol_type="genomic RNA"
 /db_xref="taxon:11103"
 /clone="JCH-5"
 /notes="isolated from chronic hepatitis patient/genotype 2a"
 2a

CDS

341. .9442
 /codon_start=1
 /product="polyprotein"
 /protein_id="BAB32877.1"
 /db_xref="GI:1312272"
 /translation="MSTNPKPQRKRNTRRRPQDKVPPGGRIYGVGVLLPRRGRPL
 GVRATKTSRSPRRORIPKDRSTCKSGKPGYPMPLVNEGLGWAGLLSPRG
 SPSCGPTDPRHSRNVKVIDITCGPADLMGYIPVCGALPGVBARALAHGVRLD
 GYNFATGNLPGFSFPLIALLSCTIIPVSAVEVRNISDSYVWTDNNSNITWQLEA
 AVLHPCGVCERKSNICWIPVSNVAPQPGALITQGLRAHIDMVNMSATLCSALH
 VGLDCGVMLAAQFTIVSPQHMFVQECNSIYPGTITGHRMAWDMNWNSTATMIL
 AVAMRVEVIIDIIISGHWGVFGLAYFSMQGAWKVVVILLIAGVADGATHTTGA
 ASNARGLASLTGSPKQNTQLINTNGSWHNRALNDSLOTGFTISALIYRNRNSS
 GCPRLSCENIEAFRIGALQYEDVNTNPEMRPYCHYFPNQCQIIPVAPQVCGPV
 YCFPSFVVVITDRGAPTYITWENDTNVFLNSTRPPLSGWFGCTWNSTGFTKTC
 GAPPKIRADFTNTDLCPTCFKRPDPATYIKCGSPWLTPKCLVDYPPYRLMHPYC
 TWNYTIFKVMYVGVYVHRLTAACNFTGRDRCNLEDSDRSQSPHLLSTEWAVLPSC
 FDLPALSTGLLHLHONIVDVOYMGSLPALTKYVVRWVWVLLFLLADARVCAW
 MLLQGAALAEKVLVHAASACNSFLYFVIFVAAWYIKGRAVLAAYSLTCLW
 PFCLLLALPOQAYADSHVQIGIQAALLILITLFTLPGYKTLRSLSWMLCYLLTL
 GEAWEPAPMQARGRGDIYWAATILCPGVFDITKWLAVLGVLPAYLLDALTSVA
 YFVRAHALLMRKTHAGGRYVQWALLALGRWSTYIYDHLTPMSDWAASGLRDLA
 VAVEPIIFSPMEKKVITWGAETAAACDILYGLPVSARLEILLGPTDGYTSKWRLL
 APTTAAOCTGRLGLAIIVSMTRCKTEQAGIOLSTVSQSLGTSISGLVWTVHG
 TGNKTLAGSRPPWOMYSAAEDGLCWSPSPGKTSLEPCTCGAVDLVLYTRNADVI
 RRGDKRGLSPRLSTLKSGSGGVLCPRGHAVGIIPRAAVCSGRVAKSIDIPVEA
 LLDVTSFSTNSSTPPAQYVQVHLAPTSKSTKVPVAYAAQYKGLVLPMSV
 AATLFGAYLSKXGHNPNIRTVRTVGTITSTYSGTKFLADGCGSGGVDIICD
 ECHAVDATTIAGTVDQAEATVGLTATATPSTVTPHPIETALQGEIEIP
 FYRAIPLSHIGRHLIFCHSKKCDLAAALRGMLNAVAYRGLDVSIIPTQGDV
 VVATDALMTGAGDVSIDCNVATQVVDLSLPTFTITQTPQDASRSRRGR
 TGRGLRTVRYVSTGERASGMDSVLCEYDAGAAHYELTPEVTVRLGAYNPFL
 PQCDHLEFAVFTGLTHIDHFLHSQTSKGFNFAYLVAYQATVCARAKAPPSSNDV
 MKCLRLKLTVPGLLYRLGVSNTVEVTLTHPTKYIATCQADLEVTMTSTWLAG
 GYLAAYATLWATGVSIIGRIHQRAVAVDPKEVLYEAFDEMEECASRAALIIEGQ
 RMAELMKIQLQOAKSOAIOPAVOASWPKVQFQWAKHWNFIISGIOYLAGLST
 LPCNPVASMFAPSAALTSTPLSTTILNILLGLMASOIAPPAGATGPPVSVGLVAA
 VGSVGLKVLVDILAGYGTIGSALVAFKIMSGEKPSMEDVNLPLGILSPCALVGV
 ICALLRRHVCEGQVQWNLHIFASIGNHVAPHYVTESDASQRTVQLLGSUTIT
 SLRLRHNVITEDTIPPCAGSLSDMDWVCTILTDPKNWLTSKLPFPKPGLPFTISQ
 KGYYGWAGTGTTPCPCGASTSGNRLGSMRITGPKTCNNIWOQTFPINCYTRGQ
 MKPADNFKTATWRVAASAEVTOHGPYSYITGLTDLNKLVPOLPSPEFESWVDGV
 QIHRFAPTPKPFRRDEVSFCVGLSVFVGSOLPCDPEPDTDLVMSLTDPSHITAEA
 ARRLARGSPPSASSASOLSPLRATCTTGKTYDMDMVEANLPMGDDVTRISES
 KVVVLDLSLMVEERDLBPSIPSEYMLPKTFPPALPAWAPDYNPLPLVESWKPDPY
 QPTTVAGCALPPPKTTPPRKRTVGLSESTIGDALQOLAIKTFGPVPPSGDSGLS
 TGAADASGRTVPDSESALSETSISSMPELEPGDPLEQVLOPPOGROGVAP
 GSDSGMSWTSCEEDSVCSPMSYWTGALITPCSPEREKLPINLPSNLSLRHNVKY
 CTTSKASIRAKKVTFRDMQVLDAYDVSVLKDIKLAASKVARSALLTLERACQITPHS
 ARSKYGFGEVRSLEGRVNHKISVWKOLLEDSQTPITPTIMAKNEVFCVDPFTGGK
 KAARLYVPLDGVRCERKALYDVTKLPQAVMGASGYQSPAPQVFEVLLKAWADKR
 EMFGYSYDFTRDTERDIERESIYQALSCAABEAGTAIHSLTERLYVGGPMLNSQ
 GOAGYVRCRACSVLTSMGNTITCVKALAAKKAAGIIVAPTMLVGGDDLVVISQ
 TREDERLAFETAMRYSAAPGDPPEYDLELITSCSNVSVVALPGQRRRYLTR
 DPTTPARAWEIVRHSVPNSWNLGNIIVQPTIWRVWMLTHFSLMAQDTLDQNLN
 FEMIGYSVSPDLDPATIERLHGLDQFSLTYSHELTRVASALKRGAPPRRA
 RARAVRASLIISGRRAICGRYLFNVAWTKLKLTPLEARLLDLSNFWTVGAGGDI
 YHSVSRAPRLLLSLLLLVGVGLLLLPAR"

BASE COUNT 1985 a 2828 c 2693 g 2185 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 9691
 Score: 15350.00 Matches: 2860
 Percent Similarity: 96.97% Conservative: 81

Best Local Similarity:	94.30%	Mismatches:	92
Query Match:	95.22%	Indels:	0
DB:	14	Gaps:	0
US-09-980-559-2 (1-3033) x AB047644 (1-9691)			
QY	1	MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGln	20
DB	341	ATGAGCACAAATCTTAAACCTCAAGAGAAAAACCAAGAAACCAACCAACCTGCCACAA	400
QY	21	AspValLysPheProGlyGlyGlnIleValGlyGlyValTyrIleuLeuProArgArg	40
DB	401	GACGTTAAGTTTCGCGCGCGCGCGGATCGTTGGCGGAGTATCTTGTTCGCGCGCAGG	460
QY	41	GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly	60
DB	461	GGCCCCAGGTTGGGTGTGCGCGCAGCAGGAAGACTTCGGAGCGGTCCACGCCACGTTGGG	520
QY	61	ArgArgGlnProIleProLysAspArgArgSerThrGlyLysSerTrpGlyLysProGly	80
DB	521	AGCGGTACGCCCATCCCTAAAGATCGGCGCTCCATCGCAAGTCTCTGGGGAACACGGA	580
QY	81	TyrProTrpProLeuTyrGlyAsnGluGlyLeuGlyTrpAlaGlyTrpIleuLeuSerPro	100
DB	581	TACCCCTGGCCCTATATGGGAATGAGGGGCTCGGCTGGCGCAGGATGGCTCTCTGCCCC	640
QY	101	ArgGlySerArgProSerTrpGlyProAsnAspProArgGluHisArgSerArgAsnValGly	120
DB	641	CGAGGTTCCCGTCCCTCTTGGGGGCGCTACTGACCCCGGCAATAGGTCGCGCAATGGGGT	700
QY	121	LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProValVal	140
DB	701	AAAGTCATCATACCATACATAACATCGCGCTTTCGCGACCTCATCGGGGTACATCCCGTCGT	760
QY	141	GlyAlaProLeuGlyGlyValAlaAlaAlaGluAlaHisGlyValArgValLeuGluAsp	160
DB	761	GGCGCCCGCTTGGTGGCGTTCGCCAGAGCTCTCGCGCAGCGGCTGAGAGTCTCTGGAGGAC	820
QY	161	GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla	180
DB	821	GGGGTCAATTTTGCAACAGGGAATTTACCTGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT	880
QY	181	LeuLeuSerCysIleThrThrProValSerAlaAlaGluValValLysAsnIleSerThrGly	200
DB	881	CTACTGTCTGTGCATCACCATTTCGGTCTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT	940
QY	201	TyrMetValThrAsnAspCysThrAsnAspSerIleThrTrpGlnLeuGlnAlaVal	220
DB	941	TACATGGTGTACTTAATGACTGTTCCTCAATAACAGCATCACCTGGCAGCTTGGGCTGCAGTC	1000
QY	221	LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTrpIle	240
DB	1001	CTCCACGTCCTCGGGTGGCTCCGCGAGAAAGTGGGCAACACATCATCGGTGCTGGATA	1060
QY	241	ProValSerProAsnValAlaValGlnArgProGlyAlaLeuThrGlnGlyLeuArgThr	260
DB	1061	CCGCTCTACCAAAAGTGGTGTCTCGGAGCCCGCGGCGCTCAGCAGGAGGCTTGGCGGG	1120
QY	261	HisIleAspMetValValMetSerAlaThrLeuCysSerAlaLeuTyrValGlyAspLeu	280
DB	1121	CACATCGACATGGTTGTGATGTCGCCACGCTTGTCTCGCTCTCCACGCTGGGGGACCTC	1180
QY	281	CysGlyGlyValMetLeuAlaAlaGlnMetPheIleValSerProGlnHisIleTrpPhe	300
DB	1181	TGCGCGGGGTGATGCTCGACCCAGACGTTTCATTTGTCTCGCGCAGCAGCAGCAGCTGTT	1240
QY	301	ValGlnAspCysAsnCysSerIleTyrProGlyThrIleThrGlyHisArgMetAlaTrp	320
DB	1241	GTGCGAGGATGCAATTTGCTCCATTTACCTTGCACCATCATCTGGACATCTGTATGATGG	1300
QY	321	AspMetMetMetAsnTrpSerProThrAlaThrMetIleLeuAlaTyrAlaMetArgVal	340
DB	1301	GACATGATGATGAACCTGGTTCGCCACCGCCACCATGATCTCTGGCGGTACGCTATGCGGTT	1360

2441 CysCysSerMetSerTyrThrThrGlyAlaLeuIleThrProCysSerProGluGlu 2460
 7661 TGCTGCTCCATGTCATATCTCTGGACCGGGCTCTAATACTCTTGTAGCCCGGAGAG 7720
 2461 GluLysLeuProIleAsnProLeuSerAsnSerLeuLeuArgTyrHisAsnLysValTyr 2480
 7721 GAAAGTTGGCAATTAACTCTCTTGGCAACTCGCTGTGGGATACCAATCAAGGTATAC 7780
 2481 CysThrThrThrLysSerAlaSerLeuArgAlaLysLysValThrPheAspArgMetGln 2500
 7781 TGTAATACATCAAGAGTGGCTCACTAAGGGCTAAAGAGTAATCTTTGATAGATGCA 7840
 2501 ValLeuAspSerTyrTyrAspSerValLeuLysAspLysLeuAlaLysLysVal 2520
 7841 GTGCTGACCGCTTATGATTCAGCTTAAAGGACATCAAGCTAGGGCTTCAAGGTC 7900
 2521 ThrAlaArgLeuLeuThrMetGluGluAlaCysGlnLeuThrProProHisSerAlaArg 2540
 7901 AGCGCAAGGCTCTCACTTATAGAGGAGCGTGGCAATTTGACCCACCCCACTCTGCAAGA 7960
 2541 SerLysTyrGlyPheGlyAlaLysGluValArgSerLeuSerGlyArgAlaValAsnHis 2560
 7961 TCCAAGTATGGGTTTGGGGCTAAGAGGTCGAGCTTGTCCGGAGGGCGGTCAACAC 8020
 2561 IleLysSerValTyrLysAspLeuLeuGluAspSerGluThrProIleProThrIle 2580
 8021 ATCAAGTCCGTGTGGAGGACCTCTTGAAGAGCTCACAACACCAATTCCTACACCATC 8080
 2581 MetAlaLysAsnGluValPheCysValAspProThrLysGlyLysLysAlaLys 2600
 8081 ATGGCCAAATAGAGTGTCTGCTGGACCCCGCCCAAGGGGGGTAAAAAACAGCTCGC 8140
 2601 LeuIleValTyrProAspLeuGlyValArgValCysGluLysMetAlaLeuTyrAspIle 2620
 8141 CTTATCGTTTACCTTGACCTCGGGCTCAGGGTCTGCGAGAGATGGCCCTTTATGATGC 8200
 2621 ThrGlnLysLeuProGlnAlaValMetGlyAlaSerTyrGlyPheGlnTyrSerProAla 2640
 8201 ACACAAAGCTTCTCAGGGGGTATGGGGCTCTTATGGCTTCCAGTACTCCCGGCT 8260
 2641 GlnArgValGluPheLeuLeuLysAlaTyrPalaGluLysLysAspProMetGlyPheSer 2660
 8261 CAGCGGTGGAGTTCTCTTGAAGCATGGCGGAAAGAGAGACCCCTATGGGTTTTCG 8320
 2661 TyrAspThrArgCysPheAspSerThrValThrGluArgAspIleArgThrGluGluSer 2680
 8321 TATGATACCGGATCTTGTACTCAACCGTCACTGAGAGACATCAGGAGTGGAGTCC 8380
 2681 IleTyrArgAlaCysSerLeuProGluGluAlaHisThrAlaLysSerLeuThrGlu 2700
 8381 ATATACAGGCTGCTCTTACCCGAGAGGCGCCGAACTGCCATACACTCGCTGACTGAG 8440
 2701 ArgLeuTyrValGlyGlyProMetPheAsnSerLysGlyGlnThrCysGlyTyrArgArg 2720
 8441 AGACTTATGTGGAGGGCCCATGTTCAACAGAGGCGCCAGTCTCGGGGTACAGGCGT 8500
 2721 CysArgAlaSerGlyValLeuThrThrSerMetGlyAsnThrIleThrCysTyrValLys 2740
 8501 TGCCGCGCCAGCGGGGTCTTACCACTAGTATGGGGAAACCAATCAGTCTATGTAAAA 8560
 2741 AlaLeuAlaAlaCysLysAlaAlaGlyIleAlaProThrMetLeuValCysGlyAsp 2760
 8561 GCCCTAGGGCTTGCAGGCTTGGGGGATATTTGGCCACCATGCTGGTATGGGGGAC 8620
 2761 AspLeuValValLysSerGluSerGlnGlyThrGluGluAspGluArgAsnLeuArgAla 2780
 8621 GACTTGGTGGTCACTCTCAGAAAGCCAGGGAGTCTGAGGAGGAGCGAGCACTGAGAGCC 8680
 2781 PheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProArgProGluTyr 2800
 8681 TTCAGGGAGGTATGACAGGATATTTGCCCCCTCTGTTGAGCCCCCGGAGCAATAT 8740
 2801 AspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaLeuGlyProGlnGly 2820

8741 GACCTGGAGCTAATAACATCTTGTCTCTCAACAGTGTCTGTGGCACTTGGCCACAGGCG 8800
 2821 ArgArgArgTyrTyrLeuThrArgAspProThrThrProIleAlaArgAlaLysGlu 2840
 8801 CCGCGGAGATATCTACTGACGAGACCCACCACTTCANTTGCCGGGCTGCTGGGAA 8860
 2841 ThrValArgHisSerProValAsnSerTrpLeuGlyAsnIleIleGlnTyrAlaProThr 2860
 8861 ACAGTTAGACACTCCCTGTCAATTCATGCTGGGAAACATCATCCAGTACGCTCCAACC 8920
 2861 IleTrpAlaArgMetValLeuMetThrHisPheSerIleLeuMetAlaGlnAspThr 2880
 8921 ATATGGTTCGATGCTCTGATGACACACTTCTTCTCCATTTCTCATGTGCCGAGACACC 8980
 2881 LeuAspGlnAsnLeuAsnPheGluMetTyrGlyAlaValTyrSerValSerProLeuAsp 2900
 8981 CTAGACCAAGCTTAATTTGAATGTACGATCGGTACTCTCCGTGAGTCTCTGGAC 9040
 2901 LeuProAlaIleIleGluArgLeuHisGlyLeuAspAlaPheSerLeuHisThrTyrThr 2920
 9041 CTCCCAAGCTAATTTGAAGGTTTACAGCGCTTTCCTCTCTGCACATACACT 9100
 2921 ProHisGluLeuThrArgValAlaSerAlaLeuArgLysLeuGlyAlaProProLeuArg 2940
 9101 CCCCACGAATCAGCGGGTGGCTTCAGCCCTCAGAAAACTTTGGGGCGCCACCCCTCAGA 9160
 2941 AlaTrpLysSerArgAlaAlaArgAlaValArgAlaSerLeuIleSerArgGlyGlyArgAla 2960
 9161 GCGTGAAGAGTCGGGGCGGTGAGTTAGGGGCTCCCTCATCTCCGCTGGGGGAGGGGG 9220
 2961 AlaValCysGlyArgTyrLeuPheAsnTrpAlaValLysThrLysLeuLysLeuThrPro 2980
 9221 GCGTTTGGGTCGGTACCTTTCACCTGGGGGTGAAGCAAGCTCAAACTCACTCCT 9280
 2981 LeuProGluAlaArgLeuLeuAspLeuSerSerTrpPheThrValGlyAlaGlyGly 3000
 9281 TTGCCGAGGAGCGCTCTCGGATTTTCCAGTTGTTTACCGTCCGGCGCGGGGGGG 9340
 3001 AspIleTyrHisSerValSerArgAlaArgProArgLeuLeuPheGlyLeuLeuLeu 3020
 9341 GACATTTATCAGAGCTGCGGTGCCGCGCCAGCCCGCTTATCTCTTACCTTACTCTTA 9400
 3021 LeuPheValGlyValGlyLeuPheLeuLeuProAlaArg 3033
 9401 CTTTCTGTAGGGGTAGGCTCTTCTTACTTCCCGCTCGA 9439

RESULT 14
 AB047644
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

AB047644 9691 bp RNA linear VRL 04-AUG-2001
 Hepatitis C virus gene for polyprotein, complete cds, clone:JCH-5.
 AB047644
 AB047644.1 GI:13122271
 Hepatitis C virus
 Hepatitis C virus
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 Hepacivirus.
 1
 Kato, T., Furusaka, A., Miyamoto, M., Date, T., Yasui, K., Hiramoto, J.,
 Nagayama, K., Tanaka, T. and Wakita, T.
 Sequence analysis of hepatitis C virus isolated from a fulminant
 hepatitis patient
 J. Med. Virol. 64 (3), 334-339 (2001)
 21316767
 11424123
 2 (bases 1 to 9691)
 Kato, T., Wakita, T. and Furusaka, A.
 Direct Submission
 Submitted (23-AUG-2000) Takanobu Kato, The Tokyo Metropolitan
 Institute for Neuroscience, Department of Microbiology, Musashidai
 2-6, Fuchu, Tokyo 183-8526, Japan (E-mail: takatomin.ac.jp,
 Tel: 81-423-25-3881 (ex. 4605), Fax: 81-423-21-8678)

Db 5441 GACAAGAGTGCTCTATAGAGGCTTTGATGAGATGAGGAATGTGCTCCAGAGCGGCT 5500
 QY LeuIleGluGlyGlnArgIleAlaGluMetLeuLysSerLysIleGlnGlyLeuLeu 1740
 Db 5501 CTCTTGAAGAGGGCGAGCGAGATAGCCGAGATGCTGAAGTCCAGATCCAAAGCTTATTG 5560
 QY GlnGlnAlaSerLysGlnAlaGlnAspIleGlnProThrValGlnAlaSerTrpProLys 1760
 Db 5561 CAGCAAGCTCTAAACAGAGCCAGGACATACACCCGCTGTGCAAGCTTCGTGGCCCAAG 5620
 QY ValGluGlnPheTrpAlaLysHisMetTrpAsnPhelIleSerGlyIleGlnTrpLeuAla 1780
 Db 5621 ATGAGCAATTCGGGCAACATATGTGAACTTCATAAGCGGCATTCAGTACCTCGCA 5680
 QY GlyLeuSerThrLeuProGlyAsnProAlaValAlaSerMetMetAlaPheSerAlaAla 1800
 Db 5681 GGACTGTCAACACTGCCAGGAAACCTGCTGTGGCTTCCATGATGGCATTCAGCGCGGCC 5740
 QY LeuThrSerProLeuSerThrSerThrThrIleLeuLeuAsnIleLeuGlyGlyTrpLeu 1820
 Db 5741 CTCACCACTCGTGTCACTAGCACCACTCTCTTAAACATTCGTGGGGGCTGGCTG 5800
 QY AlaSerGlnIleAlaProProAlaGlyAlaThrGlyPheValValSerGlyLeuValGly 1840
 Db 5801 GCGTCCCAATTCGCGCACCGCGGGGCGCACTGGCTTTGTGTGATGGCTGTGGTGGGA 5860
 QY AlaAlaValGlySerIleGlyLeuGlyLysValLeuValAspIleLeuAlaGlyTrpGly 1860
 Db 5861 GCTGCTGTTGGAGCATAGGCTTGGGTAAAGTCTGTGGACATCTCTGGCAGGGTATGGT 5920
 QY AlaGlyIleSerGlyAlaLeuValAlaPheLysIleMetSerGlyGluLysProSerMet 1880
 Db 5921 GCGGGCATTCGGGGGCGCTCGTGGCTTAAAGTCAATCATGTCTGCGGAGAGCCCTCCATG 5980
 QY GluAspValValLeuLeuProGlyIleLeuSerProGlyAlaLeuValValGlyVal 1900
 Db 5981 GAGGATGTCAATCACTGCTGCTGGGATTCGTCTCCAGGTCTCTGGTGGTGGGAGTC 6040
 QY IleCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrpMet 1920
 Db 6041 ATCTGCGCGGCCATTCGCGGCCCATGTGGAGCCGCGGGAAGCGCGGTCCAAATGGATG 6100
 QY AsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaProThrHisTrpValThr 1940
 Db 6101 AACAGCTTATCGCTTCAGAGGAAACACAGTGGCCCTACTACTACTAGTGAGC 6160
 QY GluSerAspAlaSerGlnArgValThrGlnLeuLeuGlySerLeuThrIleThrSerLeu 1960
 Db 6161 GAGTCGGATGCGTCGAGCGGTGTACCCAACTGCTTGGCTCTCTCACTATAACTAGTCTA 6220
 QY LeuArgArgLeuHisMetTrpIleThrGluAspCysProIleProCysGlyGlySerTrp 1980
 Db 6221 CTCAGGAGACTTCACAACTGGATCACTGAGGATGGCCCATGCCATGCCCGCGCTCGTG 6280
 QY LeuArgAspValTrpAspTrpValCysThrIleLeuThrAspPheLysMetTrpLeuThr 2000
 Db 6281 CTCGGCGATGTGGGCTGGCTGTGTACCATCTTAACAGACTTAAAGAACTGGCTGACC 6340
 QY SerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTrpLys 2020
 Db 6341 TCCAACTGTTCCTCAAGATGCTGGGCTCCCTCTTATCTTCTGCCCCAAAGGGGTACAAG 6400
 QY GlyValTrpAlaGlyThrGlyIleMetThrArgCysProCysGlyAlaLeuIleSer 2040
 Db 6401 GCGGTGTGGCGCGGCACTGACATCATGACACAGATGGCCCTGCGCGCGCCAACTCTCT 6460
 QY GlyAsnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTrp 2060
 Db 6461 GGCNAAGCTCGCTTGGGCTCTATGCAATCATAGGACCCCAAACTTCATGACACCTGG 6520
 QY GlnGlyThrPheProIleAsnCysTrpThrGluGlyGlnCysValProLysProAlaPro 2080
 Db 6521 CAGGGAGACCTTCTCTATCAATTGTATTACAGAGGCCAGTGTCTGCGGAAACCCGCGTTA 6580

QY 2081 AsnPheLysValAlaIleTrpArgValAlaAlaSerGluTrpAlaGluValThrGlnHis 2100
 Db 6581 AACTTCAAGACCCCATCTGGAGATGGCGGCTCAGAGTACCGGAAGTACGACGACAC 6640
 QY GlySerTrpHisTrpIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeu 2120
 Db 6641 GGATCATATGCTATATAACAGGGCTGACCACTGACACACTTAAAGTCCCTTGCACACTC 6700
 QY ProSerProGluPhePheSerTrpValAspGlyValGlnIleHisArgPheAlaProThr 2140
 Db 6701 CCCTCTCCAGAGTTTCTCTTGGGTGGAGCGAGTACAAATCCATAGTCCGCGCCACCA 6760
 QY ProLysProPhePheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal 2160
 Db 6761 CAAAGCCGTTTTCGGGATGAGGTCTGTTTCCAGCTTGGGCTCAATTCATTTGCTGTC 6820
 QY GlySerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThr 2180
 Db 6821 GGGTCTCAGCTTCCCTGTGACCTGAGCCCGACACTGAGGTAGTGTCCATGCTTAACA 6880
 QY AspProSerHisIleThrAlaGluThrAlaAlaArgArgLeuAlaArgGlySerProPro 2200
 Db 6881 GACCCATCCCATATCACGGCGGAGGCTGACGCGCGCTTTAGCGCGGGGTCAACCCCA 6940
 QY SerGluAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThr 2220
 Db 6941 TCTGAGGCAAGCTCTCAGCAGCCAGCTGTGCGGCCCATCGCTGCGAGCCACTGCACC 7000
 QY ThrHisGlyLysAlaTrpAspValAspMetValAspAlaAsnLeuPheMetGlyGlyAsp 2240
 Db 7001 ACCCAGGTAGGACCTATGATGTGACATGGTGGATGCCAACCTGTTCATGGGGGCGGC 7060
 QY ValThrArgIleGluSerGlySerLysValValValLeuAspSerLeuAspProMetVal 2260
 Db 7061 GTGATTCGATAGAGTCTGAGTCCAAAGTGGTCTGTCTGGACTCTCCCTCGACTCAATGACC 7120
 QY GluGluArgSerAspLeuGluProSerIleProSerGluTrpMetLeuProLysLysArg 2280
 Db 7121 GAGGAAGAGGGGACCTTGAGCTTCACTGATCCAGTATATGCTTCCCGAGAGAGG 7180
 QY PheProProAlaLeuProAlaTrpAlaArgProAspTrpAsnProProLeuValGluSer 2300
 Db 7181 TTCCCAACCGCTTACCGGCTTGGCGCGGCTGATTACAAACCCACCGCTGTGGATCG 7240
 QY TrpLysArgProAspTrpGlnProAlaThrValAlaGlyCysAlaLeuProProProArg 2320
 Db 7241 TGGAAAGAGCCAGATTACCAACCACTGTTGGGGGTGTGCTCTCCCGCCCGCCCAAA 7300
 QY LysThrProThrProProArgArgArgArgThrValGlyLeuSerGluAspSerIle 2340
 Db 7301 AAGACCCGACGCTCTCCCAAGGAGACCGCGGAGTGGGTCTGAGCGAGAGACCATTA 7360
 QY GlyAspAlaLeuGlnLeuAlaIleLysSerPheGlyGlnProProProSerGlyAsp 2380
 Db 7361 GGAGATGCCCTCAACAGCTGGCCATCAAGTCTTTGGCCAGCCCGCCCGCCAGCGCGAT 7420
 QY SerGlyLeuSerThrGlyAlaGlyAlaAlaAspSerGlySerGlnThrProProAspGlu 2380
 Db 7421 TCAGGCTTTTCCACGGGGCGGACCGCGCGATCTCGGCGATCGGACACCCCTCGACGAG 7480
 QY LeuAlaLeuSerGluThrGlySerIleSerSerMetProProLeuGluGlyGluLeuGly 2400
 Db 7481 TTGGCTCTTTTGGAGACAGTCTTACCTCTCATGCCCGCCCTCGAGGGGAGCGCTGGG 7540
 QY AspProAspLeuGluProGluGlnValGluProGlnProProProGlnGlyValAla 2420
 Db 7541 GACCCAGACCTGAGGCTGAGCAGGTAGAGCTTCAACCTCTCCCGAGGGGGGAGGCA 7600
 QY AlaProGlySerAspSerGlySerTrpSerThrCysSerGluGluAspAspSerValVal 2440
 Db 7601 GCTCCGGCTCGGACTCGGGTCTGTGCTTACTTGTCTCGAGGAGGATGACTCCGCTGCTG 7660

981 SerProMetGluLysValIleValTyrGluAlaGluThrAlaAlaCysGlyAspIle 1000
 3281 AGTCCCATGAGAGAAAGTCACTCGTTGGGAGCGGAGACGGCTGGTGGGGACATC 3340
 1001 LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp 1020
 3341 TTGCAGCGGACTTCCCGTGTCCGCGGACTCGGTGGGAGATCCCTTGGCCAGCTGAT 3400
 1021 GlyTyrThrSerLysGlyTyrSerLeuLeuAlaProIleThrAlaTyrAlaGlnThr 1040
 3401 GGCTACACCTCCCAAGGGGTGAAGCTTCTCGCCCATACCTTACGCCACAGACACA 3460
 1041 ArgGlyLeuLeuGlyThrIleValSerMetThrGlyArgAspLysThrGluGlnAla 1060
 3461 CGAGGTCCTTGGGCTCTAGTGGTGAGCATGACGGGGGTGACAAGACAGAACAGGCC 3520
 1061 GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly 1080
 3521 GGGGAGGTCCAAGTCTCTGTCACAGTCACTCAGTCTCTTCTCGGAACATCCATTTCGGGG 3580
 1081 ValLeuThrThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro 1100
 3581 GTCTTATGGACTGTATTACACGAGCTGGCAACAAGACACTAGCCGGCTCGCGGGCCCG 3640
 1101 ValThrGlnMetTyrSerSerAlaGluLysLeuValGlyTyrProSerProProGly 1120
 3641 GTACCGAGATGTACTCAGCGCCGAGGGGACTTGGTGGGTGGCCAGCCCTCTCTGGG 3700
 1121 ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArgAsn 1140
 3701 ACCAAATCTTTGGAGCGGTGTAGTGTGGAGCGGTGCACCTGATTATTTGGTCAGCGGAAC 3760
 1141 AlaAspValIleProAlaArgArgArgGlyAspLysArgGlyAlaLeuLeuSerProArg 1160
 3761 GCTGATGTCTATCCCGCTCGAAGACGCGGGGACAAGCGGGGAGCGCTGCTCTCCCGAGA 3820
 1161 ProLeuSerThrLeuLysGlySerSerGlyGlyProValLeuLysProArgGlyHisAla 1180
 3821 CCCCTTTGACCTTGAAGGGGTCTCTCGGGGGGACTGTGTGTTGCCCTAGGGGGCCAGCT 3880
 1181 ValGlyValPheArgAlaAlaValCysArgArgGlyValAlaLysSerIleAspPheIle 1200
 3881 GTCGGAATCTCCGGGAGCTGTGTGCTCTCGGGGTGTGGCTAAGTCCATAGATTTCATC 3940
 1201 ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro 1220
 3941 CCGGTTGAGAGCTCGACATCGTCACGGGGTCTCCACCTTTAGTGACAACAGCACACCA 4000
 1221 ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLys 1240
 4001 CCAGCTGTGCCAGACCTATCAGTGGGGTACTTGCACGGCCCTCCACTGGCAGTGGAAGA 4060
 1241 SerThrLysValProValAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnPro 1260
 4061 AGCACCAAGGTCCCGTCCGCTACGCGCCCGCGGGGTATTAAGTGTGGTGTCAATCCC 4120
 1261 SerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerLysAlaHisGlyIleAsnPro 1280
 4121 TCGGTGGTGTCCAGCTTGGGATTTGGGGCGTACTTGTCCAAAGGCACATGGCATCAACCCC 4180
 1281 AsnIleArgThrGlyValArgThrValThrThrGlyAlaProIleThrTyrSerThrTyr 1300
 4181 AACATTAGACTGGAGTCAGAACCTGTGACGACCGGGGAGCCCATTCATCTCCACGTAT 4240
 1301 GlyLysPheLeuAlaAspGlyGlyCysAlaGlyGlyAlaTyrAspIleIleCysAsp 1320
 4241 GGTAAATTCCTCGCGATGGGGGTGCGCAGCGCGCGCTATGACATCATCATATGCGAT 4300
 1321 GluCysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGlnAla 1340
 4301 GAATGCCACTCTGTGGATGCTACCACTATTCTCGGCATCGGGGAGTCTCTTGACCAAGCA 4360
 1341 GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProProGlySerValThr 1360

4361 GAGACAGCCGGGTAGGCTAAGTGTACTGGCCACGGCCACGCCCGCGGTGGGTGACA 4420
 1361 ThrProHisProAsnIleGluGluValAlaLeuGlyGlnGluGlyGluIleProPheTyr 1380
 4421 ACCCCCATCCCAATATAGAGAGGTAGCCCTCGGACAGAGGGTGAATCCCTCTCTAT 4480
 1381 GlyArgAlaIleProLeuSerTyrIleLysGlyArgHisLeuIlePheCysHisSer 1400
 4481 GGGGGGGCTTTCCCTGCTTACATCAAGGGGAGGAGCACTTGATTTCTGCCACTCA 4540
 1401 LysLysLysCysAspGluLeuAlaAlaLeuArgGlyMetGlyLeuAsnSerValAla 1420
 4541 AAGAAAAAGTGTGACAGCTCGAACCGGCCCTTCGGGGCATGGGCTTGAACGCTGGGCA 4600
 1421 TyrTyrArgGlyLeuAspValIleProThrGlnGlyAspValValValAlaAla 1440
 4601 TATTACAGAGGTGTGAGCTTCCATATACCACTCAAGGAGATGTGGTGGTCTGCTGCC 4660
 1441 ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnVal 1460
 4661 ACCGACGCCCTCATGACGGGTATCTGGAGACTTTTGACTCCGCTGATCGACTGCAACGTA 4720
 1461 AlaValThrGlnValValAspPheSerLeuAspProThrPheThrIleThrThrGlnIle 1480
 4721 CGCGTCAACCCAGCCGTAGACTTTCAGCTGGACCCCACTTCACTATAACACACAGACT 4780
 1481 ValProGlnAspAlaValSerArgSerGlnArgArgGlyArgThrGlyArgGlyArgLeu 1500
 4781 GTCCCGCAAGAGCTGTCTACAGTAGTCAGCGCGGAGGGCGACGGGTAGAGAGACTG 4840
 1501 GlyIleTyrArgTyrValSerThrGlyGluArgAlaSerGlyMetPheAspSerValVal 1520
 4841 GGCATTTATAGTATGTTTCCACTGTGTGAGCGAGCTCAGGAATGTTTGACGTGTAGTA 4900
 1521 LeuCysGluCysTyrAspAlaGlyAlaAlaTyrTyrGluLeuThrProSerGluThrThr 1540
 4901 CTCTGTGAGTGTCTACGACGAGAGCTGTGTGTATGAGCTCTCACCGAGTGGAGAGACCC 4960
 1541 ValArgLeuArgAlaTyrPheAsnThrProGlyLeuProValCysGlnAspHisLeuGlu 1560
 4961 GTCAGGCTCAGGGCGGTATTTCAACAGCTGTGGTGTGCTGTGCGCAGGACCTTGGAG 5020
 1561 PheTrpGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThr 1580
 5021 TTTTGGGAGGCGTCTTCCCGGCTCACACATAGACGCTCATTTCTCTTCCAGACA 5080
 1581 LysGlnSerGlyGluAsnPheAlaTyrLeuThrAlaTyrGlnAlaThrValCysAlaArg 1600
 5081 AAGCAGTTCGGGGGAAATTTTCGCATACTTAGTAGCTTATCAGGCCACAGTGTGGCCAGG 5140
 1601 AlaIleAlaProProProSerTrpAspValMetTrpLysCysLeuThrArgLeuLysPro 1620
 5141 GCCAAAGCGCCCCCGCTCTGGGAGCTCATGTGGAGGTGCTTGTGCTGACTCAGCTCAGCCC 5200
 1621 ThrLeuValGlyProThrProLeuLeuTyrArgLeuGlySerValThrAsnGluValThr 1640
 5201 ACGCTTGTGGCCCTACACCTCTCTGTACCGTTTGGGCTCTGTATTACCAACGAGGTCAAC 5260
 1641 LeuThrHisProValThrLysTyrIleAlaThrCysMetGlnAlaAspLeuGluValMet 1660
 5261 CTTACACACCCCGTGACAAAATACATCGCCACATGATGCAAGCTGACCTCGAGGTGATG 5320
 1661 ThrSerThrThrValLeuAlaGlyGlyValLeuAlaAlaValAlaTyrCysLeuAla 1680
 5321 ACCAGACGTGGGTCTCGGCTGGGGAGTCTTAGCAGCCGTGGCCGCTATTTGCTTAGCG 5380
 1681 ThrGlyCysValCysIleIleGlyArgLeuHisIleAsnGlnArgAlaValAlaPro 1700
 5381 ACCGGGTGTGTTCATCATTTGGCGGTTTACACATCAACACCGAGCTGTGCTCGCTCCG 5440
 1701 AspLysGluValLeuTyrGluAlaPheAspGluMetGluGluCysAlaSerArgAlaAla 1720

Db 1061 CCAGTCTCACCAACGTTGGTGTGCGCAGCCTGGCGCCCTCACGGGGGCTTGGCGAGC 1120
 QY HisLeuAspMetValValMetSerAlaThrLeuLeuCysSerAlaLeuTyrValGlyAspLeu 280
 Db 1121 CACATCGACATGCTGTTGTCGCGACACCTCTGCTCCGCTCTCTACGTTGGGGACCTC 1180
 QY CysGlyGlyValMetLeuAlaGlnMetPheLeuValSerProGlnHisHisTyrPhe 300
 Db 1181 TGTGGCGGGTGTGCTCGGCTCCAGATGTTTATGTTCTCGCGCAGCACCACCTGGTTC 1240
 QY ValGlnAspCysAsnGlySerLeuTyrProGlyThrLeuThrGlyHisArgMetAlaTyr 320
 Db 1241 GTGAGAAATGCAATGCTCCATCTACCTCGCGCCCATCATCGGGCAGCCGATATGGATG 1300
 QY AspMetMetMetAsnTyrSerProThrAlaThrMetLeuAlaTyrAlaMetArgVal 340
 Db 1301 GACATGATGATGACCTGGTGGCCACACCATCATGATCTGGCGTACGTGATCGCGTT 1360
 QY ProGluValLeuLeuAspLeuLeuSerGlyAlaHisTyrPglyValMetPheGlyLeuAla 360
 Db 1361 CCCGAGTCAATAGACATCATAGCGAGCTCACTGGGGCGTCACTGTTGGCGCTCTGG 1420
 QY TyrPheSerMetGlnGlyAlaTyrAlaLysValValValLeuLeuLeuAlaGly 380
 Db 1421 TACTTCTATGAGGAGCGTGGCGAGGTCGTTGTCTATCTCTGCTGGCGCTCTGG 1480
 QY ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr 400
 Db 1481 GTGAGCGGTACACACACGACGTGGGAGCGCTGGGGCGCCTACCATAGCTAGCTGGCC 1540
 QY SerLeuPheAspMetGlyProArgGlnLysLeuValLeuValAsnThrAsnGlySerTyr 420
 Db 1541 AGCGCTTCTCCCTGGCGCTCGCGACACATTCAGCTCATTAATACCAATGTTAGCTGG 1600
 QY HisLeuAsnArgThrAlaLeuAsnGlyAsnAspSerLeuHisThrGlyPheLeuAlaSer 440
 Db 1601 CACATCAACGCGACCGCTGAATGCAAGATTCCTCTGCACACCGGCTTCTTCACGGGC 1660
 QY LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg 460
 Db 1661 CTGTTCTACATCAATGATTCATCTGCTGGGATGTCGGAGCGCTGTCGGCGCTGTCGC 1720
 QY SerLeuGluAlaPheArgValGlyTyrPglyAlaLeuGlnTyrGluAspAsnValThrAsn 480
 Db 1721 AACATCGAGACTTCGGATAGGATGGGGCGCTCGCATACGACACATGTCACCAAT 1780
 QY ProGluAspMetArgProTyrCysTyrHisTyrProArgGlnCysGlyValValSer 500
 Db 1781 CCAGAAGATATGAGGCGCATATGCTGGCCTACCCACCAAAACAGTGTGGCGTAGTCCCC 1840
 QY AlaLysThrValCysGlyProValTyrCysPheThrProSerProValValValGlyThr 520
 Db 1841 GCAGGAGCGGTGGCGGCCGAGTACTGTTTACCCCTAGCGCGGTGTGTGGGCGACG 1900
 QY ThrAspArgLeuGlyAlaProThrTyrThrTyrPglyGluAsnGlnThrAspValPheLeu 540
 Db 1901 ACCGATAGACTTGGAGTCTACTTACAGTGGGAGAGATGAGACATGTCCTCTTA 1960
 QY LeuAsnSerThrArgProProLeuGlySerTyrPheGlyCysThrTyrPheMetAsnSer 560
 Db 1961 TTGAACAGCACCGACCGCTCGGGTCACTGTTGGCTGACGCTGGAATGAATCCACT 2020
 QY GlyTyrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer 580
 Db 2021 GGCCTCACCAAGACCTTGGCGCGCACCCCTGCGCAGCTAGAGCTGACTTCAATACCA 2080
 QY ThrAspLeuLeuCysProThrAspCysPheArgLysHisProAspThrThrTyrLeuLys 600
 Db 2081 ACGATCTGTTGGCCCCACGACTGTTTAGAAACATCTCGAAGCCACTTACATCAAA 2140
 QY CysGlySerGlyProTyrPheProArgCysLeuLeuAspTyrProTyrArgLeuTyr 620
 Db 2141 TGTGGTTCGGGGCTTGGCTCACGCCAAAGTGTCTGTTGACTTACCCCTACAGGCTCTGG 2200

QY HisTyrProCysThrValAsnTyrThrLeuPheLysIleArgMetTyrValGlyGlyVal 640
 Db 2201 CATTACCTTGCACAGTCAATTAATCTCCACTTCAAGATCAGGATGATATGGGGGAGTT 2260
 QY GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp 660
 Db 2261 GAGCACAGGCTCATGCCCGCGTCAATTTCACTCGTGGGATCGCTGCAACTTGGAGAT 2320
 QY ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGluTyrPheAlaLeuPro 680
 Db 2321 ACGGACAGAAAGTCAACAGACTCTCTGTTGCACTCCACCGAGGAAATGGGCATTTGCC 2380
 QY CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisLeuHisGlnAsnIle 700
 Db 2381 TGCTCTTCTCAGACTTGGCGCGCTTGTGCTGCTGCTTCTCCACCTCCACCAAAATATC 2440
 QY ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrLysTyrIleValArgTyr 720
 Db 2441 GTGGACGTACAATATATATGATGCGCTGTCACTGCGCTCACACAATATATGTTCCATG 2500
 QY GluTyrValIleLeuLeuPheLeuLeuAlaAspAlaArgValCysAlaCysLeuTyr 740
 Db 2501 GAGTGGGTAGTACTCTTATTTCTGCTTAGCGGAGCGCCAGGCTCTGCGCTGTTGG 2560
 QY MetLeuLeuLeuLeuGlnAlaGluAlaLeuLeuLeuValIleLeuHisAla 760
 Db 2561 ATGCTCATCTGCTGGCGCCAGCGAGCAGCAGTGGAGAGCTGCTGCTGTCGACGCT 2620
 QY AlaSerAlaAlaSerCysAsnGlyPheLeuTyrPheValIlePhePheValAlaAlaTyr 780
 Db 2621 GCGAGCGCAGCTAGCTGCAATGGCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2680
 QY TyrIleLysGlyArgValValProLeuAlaThrTyrSerLeuThrGlyLeuTyrSerPhe 800
 Db 2681 CACATCAGGAGTGGGTGGTGGCTTGGCTTATTTCCCTTACTGCGCTGTTGGCGCTTC 2740
 QY SerLeuLeuLeuLeuAlaLeuProGlnAlaTyrAlaTyrAspAlaSerValHisGly 820
 Db 2741 TGCTTACTCTCTAGACTGCGCCAGCGAGGCTTACGCTATGATGATCATCTGTCGACGGA 2800
 QY GlnIleGlyAlaAlaLeuLeuValMetIleThrLeuPheThrLeuThrProGlyTyrLys 840
 Db 2801 CAGGTGGCGCGCTTGTGCTAGTACTGATTAATCCCTCTTTTACACTCACCCCGGGTATAG 2860
 QY ThrLeuLeuSerArgPheLeuTyrPheLeuCysTyrLeuLeuThrLeuGlyGluAlaMet 860
 Db 2861 ACCCTTCTCAGCAGTCCCTGCTGGTGGTGTCTATCTCTGACCTGCGGGAACCATG 2920
 QY ValGlnGluTyrAlaProMetGlnValArgGlyGlyArgAspGlyIleIleTyrPheAla 880
 Db 2921 GTCCAGGAGTGGGACCATCCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2980
 QY ValAlaIlePheTyrProGlyValValPheAspIleThrLysThrLeuLeuAlaValLeu 900
 Db 2981 GCCACCATATTTGGCGGGCGTAGTGTGATTAACCAACCAAGTGGCTCTTAGCGGTGCTT 3040
 QY GlyProAlaTyrLeuLeuLysGlyAlaLeuThrArgValProTyrPheValArgAlaHis 920
 Db 3041 GGGCGCTGGTGTACTCTTAAGAGGTGCTTGAACGCGCTGCCATATTTCTGACAGGCCAC 3100
 QY AlaLeuLeuTyrMetCysThrMetAlaArgHisLeuAlaGlyGlyArgTyrValGlnMet 940
 Db 3101 GCTTGTGTGAGAAATGTGCACTATGTGAGGACCTCGCGGGGGGTAGGTACGTCCAGATG 3160
 QY AlaLeuLeuAlaLeuGlyArgTyrThrGlyThrTyrIleTyrAspHisLeuThrProMet 960
 Db 3161 GGGCTATTAGCCCTTGGCAGGTGGACTGGCATTACATCTATGACCACTCACCCCTATG 3220
 QY SerAspTyrAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIlePhe 980
 Db 3221 TCGGATTTGGGCTGTAGCGGCTTGGGGAGTGGCGCTGCTGTTGAGAGCTATCATCTTC 3280

Viruses: sRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.

1 Kato, T., Furusaka, A., Miyamoto, M., Date, T., Yasui, K., Hiramoto, J., Nagayama, K., Tanaka, T. and Wakita, T.
Sequence analysis of hepatitis C virus isolated from a fulminant hepatitis patient

J. Med. Virol. 64 (3), 334-339 (2001)
21316767

PUBMED
11424123

2 (bases 1 to 9674)

Kato, T., Wakita, T. and Furusaka, A.

Direct Submission

Submitted (23-AUG-2000) Takano, K., The Tokyo Metropolitan Institute for Neuroscience, Department of Microbiology; Musaeihida 2-6, Fuchu, Tokyo 183-8536, Japan (E-mail: takato@min.ac.jp, Tel: 81-423-25-3881 (ex. 4605), Fax: 81-423-21-8678)

Location/Qualifiers

1. .9674

/organism="Hepatitis C virus"

/mol_type="genomic RNA"

/db_xref="taxon:11103"

/clone="JCH-1"

/note="isolated from chronic hepatitis patient/genotype 2a"

CDS

341. 9442
/codon_start=1
/product="polyprotein"
/protein_id="BAB32873.1"
/db_xref="GI:1312264"
Translations: "MSTNPKPQKRTKNTNRPDQVKFPGGQIVGGVILPRGPRG
GVTRKASRSRSPRORIPKRRSTGKSGKPGVPLVYNEGLWAGWLLSRG
SRPSGNDPRHRSRVGKVIDLTICGFADLLGVVGVGAPLGSVALAHGVRLVED
GVNFATNLCPCFSIFLALLSCITTPVSVAQVQNTSNYMAINATNSDINSITWQLEA
AVLHPCVPCCKMGTSCWIPVSNVAVRQGLTRGRTHIDMVLSLQCSALY
VGLCGWGLMDIFVSPQHFVQECNSIYPGATITGRMAWMDMMSPPTTIL
AYVWRPEVIDIIISGAHGMVFLAYFSMQGAKVAVILLILASGDVAYTTTGSA
GRTTSLASAFSPCARNIOLINTNGSWHINRTALCNDLSLHTGPTTALFYIHKFNS
CPTERLSACRNIEDFLGALQYDDNTVPMEDMPCVHYPPKQGVGVPAGTVCCGP
YCTFSPVVGITDRIGVPIYTGNETDVFNLNSTRPSPGSGFWGCTWNSGTFTKC
GAPPCRTADPNTSDLLCPCDFRHPKPEATYICGSGPMLTPKCLVDVYPLRWHP
TVNYSTFPMYVGGVHRLMAACNFTGRDRCLNDRSQPTPLHSTTEWALIPCS
FLDLPALSTGLHKNIDVQVWYGLSPALTOYIVRWVWVLLFLLADARVCALIM
MLLLGOAEALKKLVLHAASACNGFLYFVILVAWHIKRVPVLAAYSLTGLW
PFCLLLALPOQAYADASHVQGVGALLVILFTLTPCYKILLQSLLWLCYLLTL
AETVQWASPMQARGRGGDIWAATIFCPGVVFDITKLLAVLPGFILLRGAITRVP
YFVRHALLRMCTMTRHAGRVYQMALDAGRTGTIYDHLTPMDSMAASGLRDLA
VAVEPIIFSPMEKKVIVWGAETACGDIHLGLPVSARGLREILLGPADGVTSGWKL
APITAYAOQTRGLLSTVSWTGRDTEQAGEVQVLSVTQSFLGTSISGLVWTVHG
AGNKLTAGSPVPTQWYSAGDLVGPSPGTSKLEPCTCGAVDLVLTNRNADVI PA
RRGDKRGALLSPRLSTLKGSGGVLCPRGHAVGI FRAAVCSRGVAKSIDIFIVET
LDIVTRSPFSDNSTPAPVPTQVGLHAPTSKSTKVPVAYAAQYKVLVNSV
AATLGFGATLSKAHGINPRTVGTVTTEPTITYTGKPLADGGCAGGAYDIIICD
ECHSDVATTILIGITVDQDAETAGRLTVLATATPPGSVTPPHNIEVALGQGEIP
FYGRAPLSYIKGRHLIFCHSKKKCDLALRGMGLNAVAYRGDLDVSIIPQGDV
VVVATDMLTKGTGDPDSVIDCNVAVQADFSLDPTTITOTVPODAYSRSORRG
TCGRGLGIYRVSTGERASGMPDSVLCYCDAGAAVVELSPVETVRLRAYENTPGL
FVQDHLFEAVFTGLTHIDAHFSLQTKSGENFAYLVAYQATVCAKAPPPSHDV
MWKCLTLKFTGPTLLYRLSIVNEVTLTHPVTKYIATCMQADLVETWVWAG
GVLAAYAAQCLTCVSIIGRLHINQAVAPKVEVLYEAFDEMECAKAALEEQ
RIASMLKSKIGLQGLQKQAOIQAPVQASWPKQEQFWAKHMFNFSIGIYLAGLST
LPGNPVASMWAFSAALTSPLSTSTTLLNLGGWLASQIAPPAGATFVVVGLVGA
VGSIGLGLVDIILLAGAGISGALVAFKMSGKPSMEDVINLLPGILSPGLVGV
ICAAILRRHVGGGGAQVMMNLIAFASRGNHVAPTHYVTESDASQRTQLLSLIT
SLRLRHNLITEDCPIFCAGSLRWDVWVCTILTDKFNMLTSKLPKMPGLPISCO
KXGVWAGTIGIMTRCPGANSIGNVRLASMRITGPKTCMNTWQGTFFINCYTEGQC
LPKALPAKTAIRVAASAEVATQHSYAVITGLTDLNKLPCQLSPSEFFFSWVDV
QIIRSAATPPKPFDEVSFVGLNSFVVSQLPDPEPDDEVVMSMLTDPSSHITAEA
ARRLAGSPPEASASSQSLSAFSLRATCTTHGRTYDVMWMDANLFGGQVIRIES
KVVLDSLMSNEBEGDLEPSVSEYMLPKRFPFPALPAWAPDINPLVESWKRDPY
QPPTVAGCALPPKPTTPPPRRRTVGLSESTIGDALQOLAIKSPQPPSGSDGS
TGADAAGSDRTPELALSETGSTSMPLLEGPGDPLEPEQVLEPQPPQGGAAAP
GSDSGSNSTCSEEDSVVCCSMYSWTGALITPCSPEEKLPINLSNLSLLRYHNKYV

2361	QY	SerGlyLeuSerThrGlyAlaGlyAlaAlaAaspSerClySerClnThrProProAaspGlu	2380
7387	Db	TCAGGCGCTTTCACAGGGGGCGAGCGCGCGACTCCGGCAGTCTGGTCGCCCTCGATGAG	7446
2381	QY	LeuAlaLeuSerGluThrGlySerIleSerSerMetProLeuGluGlyGlyGluLeuGly	2400
7447	Db	TTGGCTCTTTCGGAACAGGTTCTACTCTCTCATGCCCCCTCGAGGGGAGCCTGGG	7506
2401	QY	AspProAaspLeuGluProGluGlnValGluProGlnProProProGlnGlyGlyValAla	2420
7507	Db	GATCCAGACTCGAGCGCTGAGCAGGTAGAGCCCAACCCCCCCCCAGGGGGGAGGTG	7566
2421	QY	AlaProGlySerAaspSerglySerTrpSerThrCysSerGluGluAaspAaspSerValVal	2440
7567	Db	GTCTCCGGCTCGGACTCGGGGTCTCGTCTACTTGTCTCGAGGAGAATGTCTCCGTCGT	7626
2441	QY	CysCysSerMetSerTyrSerTrpThrGlyAlaLeuIleThrProCysSerProGluGlu	2460
7627	Db	TTCTGTCTCCATTTCATCTCTGAGCGGGGTCTAATACTCTTGTATTCCTCCGAAAAA	7686
2461	QY	GluLysLeuProIleAenProLeuSerAenSerLeuLeuArgTyrHisAasnLysValTyr	2480
7687	Db	GAAAGTTGCCATTTTAACCCCTTAAGCAACTCCCTCTGTTCGATACGACAAAGGTGTAC	7746
2481	QY	CysThrThrThrylsSerAlaSerLeuArgAlaLysLysValThrPheAaspArgMetGln	2500
7747	Db	TGTACTACATCAAGAGAGCCCTCACTAGGGGCTTAAAGGTAACTTTTGTATAGGATGCAA	7806
2501	QY	ValLeuAaspSerTyrTyrAaspSerValLeuLysAaspIleLysLeuAlaAlaSerLysVal	2520
7807	Db	GTGCTCGACGCTATTATGACTCAGTCTTAAAGGACATCAAGCTAGCGGCTTCCAAAGTC	7866
2521	QY	ThrAlaArgLeuLeuThrMetGluGluAlaCysGlnLeuLeuThrProProHisSerAlaArg	2540
7867	Db	AGCGCAAGGCTCTCACTTGGAGGAGGGGTGCCAATTGACTCCACCCCATTCGCAAGA	7926
2541	QY	SerLysTyrGlyPheGlyAlaLysGluValArgSerLeuSerClyArgAlaValAenHis	2560
7927	Db	TCCAGATAGGGTTTGGGGCTTAAGGAGTTTCGAGGCTTGTCCGGAGAGGCCGTAAACCAC	7986
2561	QY	IleLysSerValTrpLysAaspLeuLeuGluAaspSerGluThrProIleProThrThrIle	2580
7987	Db	ATCNAAGTCGTGTGGAGGACCTCTCGGAAGACTCACAACACCAATTCCTCCACGACATC	8046
2581	QY	MetAlaLysAasnGluValPheCysValAaspProThrLysGlyGlyLysLysAlaAlaArg	2600
8047	Db	ATGGCCAAAATAGAGTGTCTCGTGGAGCCCAACAAGGGGGGTAAAGAAAGCAGCTGC	8106
2601	QY	LeuIleValTyrProAspIleGlyValArgValCysGluLysMetAlaLeuTyrAspIle	2620
8107	Db	CTTATCTGTTTACCTCTGACCTCGGTGTACAGAGTTTCGAGAAGATGGCCCTTTATGATGC	8166
2621	QY	ThrGlnLysLeuProGlnAlaValMetGlyAlaSerTyrGlyPheGlnTyrSerProAla	2640
8167	Db	ACACAAAAGCTTCTCAGCGAGTGTGGGGGCTCTTATGGCTTCAGTACTCCCCCGCT	8226
2641	QY	GlnArgValGluPheLeuLeuLysAlaTrpAlaGluLysLysAaspProMetGlyPheSer	2660
8227	Db	CAACGGGTGAGTTTCTCTTGAAGGATGCGCGGAAAAAAAAGACCCCTATGGGGTTTTCG	8286
2661	QY	TyrAaspThrArgCysPheAaspSerThrValThrGluArgAspIleArgThrGluGluSer	2680
8287	Db	TATGATACCCGATGCTTTTGACTCAACCGTCACTGAGAGAGACATCAGAACTCAGAGAGTCC	8346
2681	QY	IleTyrArgAlaCysSerLeuProGluGluAlaHisThrAlaIleHisSerLeuThrGlu	2700
8347	Db	ATATACCGGCTTGTTCCTTGGCCGAGGAGGCCACACTGCCATACATCTCGTGTACTGAG	8406
2701	QY	ArgLeuTyrValGlyClyProMetPheAenSerLysGlyGlnThrCysGlyTyrArgArg	2720
8407	Db	AGACTTTACGTGGGAGGGCCATGTTTAAACAGCAAGGCGCAGGCTTCGGGTACAGGGGT	8466

2721	Qy	CysArgAlaSerGlyValLeuThrThrSerMetGlyAsnThrIleThrCysThrValIys	2740
8467	Db	TGCGCGCCAGCGGGGTGCTACCACTAGCATAGCATGGGNAACACCATCAGTGTATGTGAAA	8526
2741	Qy	AlaLeuAlaAlaCysLysAlaAlaGlyIleIleAlaProThrMetLeuValCysGlyAsp	2760
8527	Db	GCCTTAGCGGCTGTAAAGCTGCGGGATAGTTGGCCCACTGCTGGTATGCGGGAC	8586
2761	Qy	AspLeuValValIleSerGlnSerGlnGlyThrGluGluAspGluArgAsnLeuArgAla	2780
8587	Db	GACTTGGTTGTTCATCTCAGAAAGCCAGGGGACTGAGGAGGACGACGCGAAACCTGAGAGCC	8646
2781	Qy	PheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProArgProGluTyr	2800
8647	Db	TTACGAGAGGCTATGACCAAGTACTCTGCCCTCCTCGTGAACCTCCACAGACCGGAATAT	8706
2801	Qy	AspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaLeuGlyProGlnGly	2820
8707	Db	GACCTGGAGCTGATAAACATCTTGTCTCAATGTCTGTGCATCTGGGCCCACAAGGC	8766
2821	Qy	ArgArgArgTyrTyrLeuThrArgAspProThrProIleAlaArgAlaAlaIleProGlu	2840
8767	Db	CGCGCAGACTACTACCTGACAGACACCTACCACTCCAACTCGCCGGCTGCTGGGAA	8826
2841	Qy	ThrValArgHisSerProValAsnSerTrpLeuGlyAsnIleIleGlnTyrAlaProThr	2860
8827	Db	ACAGTTAGACACTCCCTCTGTCAAATTCATGGCTGGGAAACATCATCCAGTAGGCTCAAAC	8886
2861	Qy	IleTrpAlaArgMetValLeuMetThrHisPhePheSerIleLeuMetAlaGlnAspThr	2880
8887	Db	GTATGGGTTGCATGGTCTCATGACACACTTCTCTCATTTCTCATGGCCCAAGACACC	8946
2881	Qy	LeuAspGlnAsnLeuAsnPheGluMetTyrGlyAlaValTyrSerValSerProLeuAsp	2900
8947	Db	CTGGACCAAGAACCTCAACTTTGAAATGTACGGATCAGTGACTCTGTGAGTCTTTTGGAC	9006
2901	Qy	LeuProAlaIleIleGluArgLeuHisGlyLeuAspAlaPheSerLeuHisThrTyrThr	2920
9007	Db	CTCCAGGCATTAATTGAAAGTTACAGGGCTTGACGCTTTCTCTCTGCACACATACTCT	9066
2921	Qy	ProHisGluLeuThrArgValAlaSerAlaLeuArgLysLeuGlyAlaProProLeuArg	2940
9067	Db	CCCCACGAACCTGACCGGGTGCTGACGCCCTCTAGAAACCTTGGGGCGCCACCCCTCAGA	9126
2941	Qy	AlaTrpLysSerArgAlaArgAlaValArgAlaSerLeuIleSerArgGlyGlyArgAla	2960
9127	Db	GGTGGAGAGTCGGCGCGTCAGTTAGGCGCTCCCTCATCTCCCATGGAGGAGGGCG	9186
2961	Qy	AlaValCysGlyArgTyrLeuPheAsnTrpAlaValLysThrLysLeuLysLeuThrPro	2980
9187	Db	GCCGTTTTCGGTTCGTATCTTCAACTGGGCGGTGAAGACCAAGCTCAAACTCACTCCA	9246
2981	Qy	LeuProGluAlaArgLeuLeuAspLeuSerTrpPheThrValGlyAlaGlyClyGly	3000
9247	Db	TTGCGGAAGCACGCCACCTGGATTATTCAGTTTGGTTTTACTGTGGCGCGCGGGGGC	9306
3001	Qy	AspIleTyrHisSerValSerArgAlaArgProArgLeuLeuLeuPheGlyLeuLeuLeu	3020
9307	Db	CACATTTATCACGGGTGCGGTGCCGACCCCGCTTATTACTCTTGGGCTACTCTCTA	9366
3021	Qy	LeuPheValGlyValGlyLeuPheLeuLeuProAlaArg	3033
9367	Db	CTTTTCGTAGGGGTAGGCTCTTCTACTCTCCCGCTCGG	9405

RESULT 13
AB047640
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AB047640 9674 bp RNA linear VRL 04-AUG-2001
Hepatitis C virus gene for polyprotein, complete cds, clone:JCH-1.

AB047640
AB047640.1 GI:13122263

KEYWORDS
SOURCE
ORGANISM
Hepatitis C virus
Hepatitis C virus

Db	5167	ACGCTCGTGGTCCACACCTCTCTCTGATCCGTTGGGCTCCGTTGCCAACGAGATCACC	5226	Db	6247	CTCCGTGATGTGGGATGGGTTGGCCACCATCTTAACACAGACTTTAAAAATTTGGCTGACC	6306
Qy	1641	LeuThrHisProValThrLysTyrIleAlaThrCysMetGlnAlaSerLeuGluValMet	1660	Qy	2001	SerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTyrLys	2020
Db	5227	CTCACACACCTCTGCAAAAGTATCATCGCCACATGATGCAAGCTGACCTTGAGGTCATG	5286	Db	6307	TCCAATTTGTTTCCAAAGATGCCCGCTCCCTCTTATCTCTTGCCAGAGGGGTACAAG	6366
Qy	1661	ThrSerThrTrpValLeuAlaGlyGlyValLeuAlaAlaValAlaAlaTyrCysLeuAla	1680	Qy	2021	GlyValTrpAlaGlyThrGlyIleMetThrThrArgCysProCysGlyAlaAlaSer	2040
Db	5287	ACTAGCACGTGGGTTCTAGCTGGGGAGTCTTGGCAGCCGCTCCCGGGTATTGCTTGGCG	5346	Db	6367	GGCGTGTGGCGCGGCATCGTATCATGACCAACGATGCTCTTGGCGGCCAACATCTCT	6426
Qy	1681	ThrGlyCysValCysIleIleGlyArgLeuHisIleAsnGlnArgAlaValAlaPro	1700	Qy	2041	GlyAsnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTrp	2060
Db	5347	ACCGGTGTCTTATCATCGCGCTTTCACATCAATCAGCAGCGCTGTTGGCGG	5406	Db	6427	GGTAATGTCCGCTGGCTCTATGAGATCAGAGGCTTAAACCTGATGACACTTGG	6486
Qy	1701	AspLysGluValLeuTyrGluAlaPheAspGluMetGluCysAlaSerArgAlaAla	1720	Qy	2061	GlnGlyThrPheProIleAsnCysTyrThrGluGlyGlnCysValProLysProAlaPro	2080
Db	5407	GACAAGAGGTCCTCTATGAGGCTTTTGATGAGATGCTGAAGTCCAGATCCAAAGCTTATTG	5466	Db	6487	CAGGGAGCTTCTCTATCAATTTGTACACGGAGGCGAGTGGTGGCGAAACCGCGCCA	6546
Qy	1721	LeuIleGluGlnGlyGlnArgIleAlaGluMetLeuLysSerLysIleGlnGlyLeuLeu	1740	Qy	2081	AsnPheLysValAlaIleTrpArgValAlaAlaSerGluTyrAlaGluValThrGlnHis	2100
Db	5467	CTGTTGAAGAGGGGACCGGATAGCCGAGATGCTGAAGTCCAGATCCAAAGCTTATTG	5526	Db	6547	AACTTCAAGGTCGCCATCTGGAGGTTGGCGCTCAGAGTACGCGAGGTGACCGACAC	6606
Qy	1741	GlnGlnAlaSerLysGlnAlaGlnAspIleGlnProThrValGlnAlaSerTrpProLys	1760	Qy	2101	GlySerTyrHisTyrIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeu	2120
Db	5527	CAGCAAGCTTCTAAACAGGCCAGGACATACACCCGCTGTGCAGGCTTCATGGCCCCAAG	5586	Db	6607	GGATCATCTCTTATATAACAGGATTCACCATCTGAACTTTGAAAGTTCCCTGCCCACTA	6666
Qy	1761	ValGluGlnPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeuAla	1780	Qy	2121	ProSerProGluPhePheSerTrpValAspGlyValGlnIleHisArgPheAlaProThr	2140
Db	5587	GTAGAACAAATTTGGGGTAAACACATGTGAACCTTCATCAGCGGCATTCATACCTCGCA	5646	Db	6667	CTTCTCCAGAGTCTTCTCTCTGGGTGGATGGGTGCCAATCCATAGGTTCGCTCCACA	6726
Qy	1781	GlyLeuSerThrLeuProGlyAsnProAlaValAlaSerMetMetAlaPheSerAlaAla	1800	Qy	2141	ProLysProPhePheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal	2160
Db	5647	GGACTGTCAACCTGCCAGGAAACCCGCTGTGGCTTCCATGATGGCAATTCAGTCCGCC	5706	Db	6727	CCAAAGCCGTTTTTCCGGGATGAGGTCTCGTTCTGGTGTGGACTCAATTCATTTGTGTC	6786
Qy	1801	LeuThrSerProLeuSerThrSerThrThrIleLeuLeuAsnIleLeuGlyIleTrpLeu	1820	Qy	2161	GlySerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThr	2180
Db	5707	CTCACAGTCTCTGTCACTAGCACCATCTCTCTCAACATTTTGGGAGGCTGGCTG	5766	Db	6787	GGGTCTCAGCTTCTTGTGACCTTGACCCCGCCGACACTGACGTACTGACGTCCATGCTAACA	6846
Qy	1821	AlaSerGlnIleAlaProProAlaGlyAlaThrGlyPheValValSerGlyLeuValGly	1840	Qy	2181	AspProSerHisIleThrAlaGluThrAlaAlaArgArgLeuAlaArgGlySerProPro	2200
Db	5767	GGGTCCCAAAATCGCGCCACCGCAGGGGGCCACCGGCTTCGTGCTGCTGAGTCCGCTGG	5826	Db	6847	GATCCATCCCATATACCGCGGAGACCGCAGCGCGGCTTGGCGGGGGGTCACTTCCA	6906
Qy	1841	AlaAlaValGlySerIleGlyLeuGlyLysValLeuValAspIleLeuAlaGlyTyrGly	1860	Qy	2201	SerGluAlaSerSerAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThr	2220
Db	5827	GCAGCTGTAGCAGCATAGGCTCGGCAAGTGTGTTGGACATCTCTGGCAGGGTATGTT	5886	Db	6907	TCCGAGGCAAGCTCTCTCAGCGAGCCAGCTATCAGCGCATCTGCTGCGAGCCACTGCACC	6966
Qy	1861	AlaGlyIleSerGlyAlaLeuValAlaPheLysIleMetSerGlyGlyLysProSerMet	1880	Qy	2221	ThrHisGlyLysAlaTyrAspValAspMetValAspAlaAsnLeuPheMetGlyGlyAsp	2240
Db	5887	CGGGGCATTTGGGGGGCCCTGTCGCATTTAAGATCATGCTGCGGAGAGCCCTCCATG	5946	Db	6967	ACCCATGGCAAGACTTATGATGTGGACATGTTGGATGCCACGCTGTTTCATGGGGGGCGAT	7026
Qy	1881	GluAspValValAsnLeuLeuProGlyIleLeuSerProGlyAlaLeuValValGlyVal	1900	Qy	2241	ValThrArgIleGluSerGlySerLysValValValLeuAspSerLeuAspProMetVal	2260
Db	5947	GAGGATGTTGTCAATCTGCTCGCTGGGATTTGTCTCGGGTGCCTTAGTGGTGGAGTC	6006	Db	7027	GTGACTCGGATAGAGTCTGAGTCCAAAGTGGTCTCTCTGAGCTCTCTGAGCCCAATGGCC	7086
Qy	1901	IleCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrpMet	1920	Qy	2261	GluGluArgSerAspLeuGluProSerIleProSerGluTyrMetLeuProLysLysArg	2280
Db	6007	ATCTGCGGGCCATTTCTACGCGCCACGTGGGACCGGGGGAGGTGGGTCCAATGGATG	6066	Db	7087	GNAGAGAGAGAGCCCTCGAGCCCTCGATACCATCCGAGTATATGCTCTCCAGGAACAGG	7146
Qy	1921	AsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaProThrHisTyrValThr	1940	Qy	2281	PheProProAlaLeuProAlaTrpAlaArgProAspTyrAsnProProLeuValGluSer	2300
Db	6067	AACAGGCTCATGCTTGTCTCCAGAGAAACCAACCGCTCGCCCTACCCACTACGTGAGC	6126	Db	7147	TTCCCGCCAGCTTACCGGCTCGGCAACCGGCTTGAATTACAAACCCACCGCTTGTGGAGTCG	7206
Qy	1941	GluSerAspAlaSerGlnArgValThrGlnLeuLeuGlySerLeuThrIleThrSerLeu	1960	Qy	2301	TrpLysArgProAspTyrGlnProAlaThrValAlaGlyCysAlaLeuProProArg	2320
Db	6127	GAGTCGATGCGTCCGACGCTGTGACCCCACTGCTTGGATCCCTTACTATAACTAGTCTA	6186	Db	7207	TGGAAGAGGCCAGATTACCAACCGGCCACTGTTGGGGCTGTGCTCTCCCCCCCCCTAAG	7266
Qy	1961	LeuArgArgLeuHisAsnTrpIleThrGluAspCysProIleProCysGlyGlySerTrp	1980	Qy	2321	LysThrProThrProProProArgArgArgThrValGlyLeuSerGluAspSerIle	2340
Db	6187	CTCAGGAGACTTCCAAATTTGATTTACTGAGGACTGCCCACTCCATTCGCGCGGTTCGTGG	6246	Db	7267	AAGSCCCGAGCGCCCCCCCCCAAGAGAGCGCGGACAGTGGGTCTGAGCGAGAGCACCATA	7326
Qy	1981	LeuArgAspValTrpAspTrpValCysThrIleLeuThrAspPheLysAsnTrpLeuThr	2000	Qy	2341	GlyAspAlaLeuGlnGlnLeuAlaIleLysSerPheGlyGlnProProProSerGlyAsp	2360
				Db	7327	GGAGACGCCCTCCAACAGCTGGCCATCAAGTCTCTGGCGCAGCCCCCCCCCAAGCGCGAT	7386

QY	901	GlyProAlaTyrLeuLeuLysGlyAlaLeuThrArgValProTyrPheValArgAlaHis	920	QY	1261	SerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerLysAlaHisGlyIleAsnPro	1280
DB	3007	GGGCTGCTTACCTCTTAAGAGATGCTTCACACGGTGCCTTATTTCTGTCAGAGCTCAC	3066	DB	4087	TCGGTGGCTGCCACCTTGGGGTTTGGGGCTACTGTCACAGGCACATGSCATCAATCCC	4146
QY	921	AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyGlyArgTyrValGlnMet	940	QY	1281	AsnIleArgThrGlyValArgThrValThrGlyAlaProIleThrTyrSerThrTyr	1300
DB	3067	GCTCTGCTGAGGATGTCACCGTGTGAGGAACCTCGCGGGGGTAAGTATGTCACAGATG	3126	DB	4147	AACATCAGGACTGGGGTCAGGACTGTGACACCGGGCGGCCATCATACTCCACGTAT	4206
QY	941	AlaLeuLeuAlaLeuGlyArgTyrThrGlyThrTyrIleTyrAspHisIleThrProMet	960	QY	1301	GlyLysPheLeuAlaAspGlyGlyCysAlaGlyGlyValaTyrAspIleIleCysAsp	1320
DB	3127	GGCTATTAGCCCTTGGCAGGTGACCGGCATTTACATCTATGACCACTTCACACCTATG	3186	DB	4207	GGCAAAATTCCTCCCATGCGGGCTCGCAGGGCGGCCCTATGACATCATATGCGAC	4266
QY	961	SerAspTrpAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIleIlePhe	980	QY	1321	GluCysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGlnAla	1340
DB	3187	TCGGACTGGGCTGTACGGCTCGGGACTTGGGGTCTGCTGGAGCCCATCATCTTC	3246	DB	4267	GAATGCCACGCTGTGATTTCTACCACTTCGCGCATTTGGACAGTTCTTGACCAAGCA	4326
QY	981	SerProMetGluLysLysValIleValTrpGlyAlaGluThrAlaAlaCysGlyAspIle	1000	QY	1341	GluThrAlaGlyValArgIleThrValLeuAlaThrAlaThrProProGlySerValThr	1360
DB	3247	AGTCCGATGAGAGAAAGTCAATCGCTGGGAGCAGAGACAGCGCATCGCGGATATC	3306	DB	4327	GAGACAGCGGGGTACAGGCTAACTGTACTGGCCACCGCTACGCCCGCGGTCGGTGACG	4386
QY	1001	LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp	1020	QY	1361	ThrProHisProAsnIleGluGluValAlaLeuGlyGlnGluGlyGluIleProPheTyr	1380
DB	3307	NTGACCGGCTCCCGGTGCTCGGCCGACTTGGTGGGAGATCTCTTGGCCGAGCTGAT	3366	DB	4387	ACCCCCACCCCAATATAGAGAGGTAGCCCTCGGGCAGAGGGTGAGATCTCCCTCTCAT	4446
QY	1021	GlyTyrThrSerLysGlyTrpSerLeuLeuAlaProIleThrAlaTyrAlaGlnGlnThr	1040	QY	1381	GlyArgAlaIleProLeuSerTyrIleLysGlyGlyArgHisLeuIlePheCysHisSer	1400
DB	3367	GGCTACACCTCCAAAGGGTGGAGCTTCTCGCCCCCATCACTGTTACGCTCAGCAGACA	3426	DB	4447	GGGAGGGCATTTCCCTTGTCTCATCATCAAGGGGAGGAGACACTTGATTTTCTGCCACTCA	4506
QY	1041	ArgGlyLeuLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGluGlnAla	1060	QY	1401	LysLysLysCysAspGluLeuAlaAlaLeuArgGlyMetGlyLeuAsnSerValAla	1420
DB	3427	CGAGGCTCTTGGCGCTATAGTGTGAGCATGACGGGGCGGACAGACAGAACAGGCC	3486	DB	4507	AAGAAAAAGTGTACAGCTCGCGCGGCCCTTCGGGGCATGGGCTGGAACGTGTGGCA	4566
QY	1061	GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly	1080	QY	1421	TyrTyrArgGlyLeuAspValSerValIleProThrGlnGlyAspValValValAla	1440
DB	3487	GGGGAATCAAGTCTCTCCACGGTCACCTCTCTCTCGGAACAACCATTTTCGGGG	3546	DB	4567	TATTACAGAGGGTGTAGAGCTCTCCGPAATACCAACTCAAGGAGATGTAGTGGTCTTCGCC	4626
QY	1081	ValLeuTrpThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro	1100	QY	1441	ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnVal	1460
DB	3547	GTTTTGTGACTGTTTACACGAGCTGGCAACAGACTCTAGCCGCTCAGCGGGCCCG	3606	DB	4627	ACCGAGCCCTCATGACGGGGTATCTACTGGAGACTTTTGACTCTGTGATCGATTTGCAAGCTA	4686
QY	1101	ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTrpProSerProProGly	1120	QY	1461	AlaValThrGlnValValAspPheSerLeuAspProThrPheThrIleThrThrGlnIle	1480
DB	3607	GTACGACAGATGTACTCGAGTGGCGGGGGACTTGGTAGGGTGGCCACCTCTCCCGGG	3666	DB	4687	CGGCTCACTCAAGTCTGAGACTTCAGTTTGGACCCCACTTACCATAACACACAGACT	4746
QY	1121	ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArgAsn	1140	QY	1481	ValProGlnAspAlaValSerArgSerGlnArgArgGlyArgThrGlyArgGlyArgLeu	1500
DB	3667	ACCAAAATCTTTGAGCGGTGTGCTGTGGAGCGGTGACCTGTACTGTGTGTCAGCGGAAC	3726	DB	4747	GTCCCTCAAGACCGCTCTCGCGTAGCCGCGGAGCGGTACAGGTAGGGGGAGACTG	4806
QY	1141	AlaAspValIleProAlaArgArgGlyAspLysArgGlyValAlaLeuLeuSerProArg	1160	QY	1501	GlyIleTyrArgTyrValSerThrGlyGluArgAlaSerGlyMetPheAspSerValVal	1520
DB	3727	GCTGATGTATCCGGCTCGAGACCGGGGACAGCGGGGAGCGTTACTCTCCCGAGA	3786	DB	4807	GGCATTTATAGGTATGTTTCCACTGGTAGCGGAGCTCAGGAATGTTTGACAGTGTGGTG	4866
QY	1161	ProLeuSerThrLeuLysGlySerSerGlyProValLeuCysProArgGlyHisAla	1180	QY	1521	LeuCysGluCysTyrAspAlaGlyAlaAlaTyrTyrGluLeuThrProSerGluThrThr	1540
DB	3787	CCTCTTTCGACCTTGAAGGGTCTCTCGGGGGACCGGTGCTTTGCCCTAGGGGCCACGCT	3846	DB	4867	CTCTGTAGTGTCTACGACGCGGGTGTGATGTATGAGTCTCACACGACGAGACTTACC	4926
QY	1181	ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle	1200	QY	1541	ValArgLeuArgAlaTyrPheAsnThrProGlyLeuProValCysGlnAspHisLeuGlu	1560
DB	3847	GTCCGGAATTTCCGGGACGCGGTGCTCTCGGGGGGTGGCCAAAGTCCATAGATTTTATC	3906	DB	4927	GTGAGCTTAGAGCGGTATTTCAACACGCCAGCGCTTGTGTGTGCCAAGACCATCTTGAG	4986
QY	1201	ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro	1220	QY	1561	PheTrpGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThr	1580
DB	3907	CCGCTTGAGACCTCGACATCTGTCAGCGGTCCCCCACCCTTCAGTGACAAACACACCA	3966	DB	4987	TTTTGGGAGGAGTTTTCACCGGCTCACACATAGATGCCCATCTTCTTTTCCCAACA	5046
QY	1221	ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLys	1240	QY	1581	LysGlnSerGlyGluAsnPheAlaTyrLeuThrAlaTyrGlnAlaThrValCysAlaArg	1600
DB	3967	CCAGCTGTGCCCCAGACCTACAGGTCGGGTACTTGCACGCCCGGACTGCGCGGGAAA	4026	DB	5047	AAGCAATCGGGGAGAAATTTGTCATCTTGGTAGCTATCAAGCCACAGTGTGCGCTAGA	5106
QY	1241	SerThrLysValProValAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnPro	1260	QY	1601	AlaLysAlaProProProSerTrpAspValMetTrpLysCysLeuThrArgLeuLysPro	1620
DB	4027	AGTACCAAAAGTCCCGTCCGCTACGCGCTCAGGGGTACAAAGTGTAGTGTAAATCCC	4086	DB	5107	GCCAAAGGCCCTTCCCCGCTCGGAGCGTCATGTGGAGTGTGTGACCCGACTTAAGCCC	5166
				QY	1621	ThrLeuValGlyProThrProLeuLeuTyrArgLeuGlySerValThrAsnGluValThr	1640

Db	787	GGGGTTAAATATGCAACAGGAACTTACCTGGTGTCTCTTTCTATTTCTTGTGCGCC	846	Db	1867	ACCAACAAATTTGGAGTGCCTTACACATGCGGGGAGAAATGACAGATGCTCTTTATA	1926
Qy	181	LeuLeuSerCysIleThrThrProValSerAlaAlaGluValIysAsnIleSerThrGly	200	Qy	541	LeuAsnSerThrArgProLeuGlySerTrpPheGlyCysThrTrpMetAsnSerSer	560
Db	847	CTACTGCTTCGCTGCTACCGTTCGGTCTCTGCTGCTCCAGTGAAGAACACCACTGAGACC	906	Db	1927	TTGAACAGCACCCGACACCCGCGGGGTCACTGGTTCGGTGCACGTGGATGATTCCTACT	1986
Qy	201	TyrMetValThrAsnAspCysThrAsnAspSerIleThrTrpGlnLeuGlnAlaAlaVal	220	Qy	561	GlyTyrThrLysThrCysGlyAlaProProCysAsrGthrArgAlaAspPheAsnAlaSer	580
Db	907	TACATGGTGAATATGACTGTTCAATGACAGCATACCTGGCAGCTTCAGGCGCGGCTC	966	Db	1987	GGTTTCACCAAGACTTGTGGCGCACCCACTTGGCGCATTTAGGCGCGACTTCAATGCCAGC	2046
Qy	221	LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTrpIle	240	Qy	581	ThrAspLeuLeuCysProThrAspCysPheAsgLysHisProAspThrThrTrpLeuLys	600
Db	967	CTCCAGCTCCCGGGTGGTGGTGGAGCGGAAATACATCACGGTCTGGATA	1026	Db	2047	ACGGACCTCTTGGCCCCACGACTGTTTTAGAGACATCTCTGAAGCCACTTACATCAAA	2106
Qy	241	ProValSerProAsnValAlaValGlnArgProGlyAlaLeuThrGlnGlyLeuArgThr	260	Qy	601	CysGlySerGlyProTrpLeuThrProArgCysLeuIleAspTyrProTyrArgLeuTrp	620
Db	1027	CCGGTTTACACAAACGTGGCGTGCAGCAGCCCGGCGCTTCACGAGGGCTTGGCGAGC	1086	Db	2107	TGTGGTCTTGGGCCCTGGCTCAGCCCAAGGTGCTGGTGCAGTACCCCTACAGGCTCTGG	2166
Qy	261	HisIleAspMetValValMetSerAlaThrLeuCysSerAlaLeuTyrValGlyAspLeu	280	Qy	621	HisTyrProCysThrValAsnTyrThrIlePheLysIleArgMetTyrValGlyGlyVal	640
Db	1087	CACATCGACATGGTGTGATGTCGCCACCGCTCTGCTCCGCCCTTTACGTGGGGACCTC	1146	Db	2167	CATTACCCCTGCACAGTCAATTCATCTTTAAGATAAGGATGATGTGGGGGGGTT	2226
Qy	281	CysGlyGlyValMetLeuAlaAlaGlnMetPheIleValSerProGlnHisHisTrpPhe	300	Qy	641	GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp	660
Db	1147	TGCGCGGGGTGGTCTCGCAGCCACCTGTTTCATTGCTCTCGCGCGCGCCACTGGTTT	1206	Db	2227	GAGCACAGGCTCACGCGCGCTGCAATTTCTCTCGTGGGGATGCTTGCACCTTGGAGGAC	2286
Qy	301	ValGlnAspCysAsnCysSerIleTyrProGlyThrIleThrGlyHisArgMetAlaTrp	320	Qy	661	ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGluTrpAlaIleLeuPro	680
Db	1207	GTGCAGGAATGCAATGCTCATCTATCTGGTGCCATCATCTGCACACCGTATGGCATGG	1266	Db	2287	AGGACAGGAGTCACTGACTCTCTGTTGTCATCTCCACACCGAATGGGCCATCTTCCC	2346
Qy	321	AspMetMetMetAsnTrpSerProThrAlaThrMetIleLeuAlaTyrAlaMetArgVal	340	Qy	681	CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisLeuHisGlnAsnIle	700
Db	1267	GACATGATGATGAACCTGGTGGCCACACCACTGCTTGGGCTACGCAATGCGCGTT	1326	Db	2347	TGTTCTTACTCAGACTTGCCTGCTTGCAGCTGGCTTCTCCATCTCCACCAAAATATC	2406
Qy	341	ProGluValIleIleAspIleIleSerGlyAlaHisTrpGlyValMetPheGlyLeuAla	360	Qy	701	ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrIlystYrIleValArgTrp	720
Db	1327	CCCGAGGTCAATATAGACATCATTTAGTGGGGCTCCTGGGGCTGTTGGCGCTAGCC	1386	Db	2407	GTGAGCGTCAATATACATATGATGGCCTGACACCTGCCCTCACAANAATATGCTTCATGG	2466
Qy	361	TyrPheSerMetGlnGlyAlaTrpAlaLysValValIleLeuLeuLeuAlaAlaGly	380	Qy	721	GluTrpValIleLeuLeuPheLeuLeuAlaAspAlaArgValCysAlaCysLeuTrp	740
Db	1387	TACTTCTATGACAGGAGCGTGGCGGAGTGTCCGGCGGTATCCACCAATGGCAGTGG	1446	Db	2467	GAGTGGGTAGTACTTATTTCTCTTACCTAGCGAGCGCCAGAGGTGGCGCTGCTTGGTGG	2526
Qy	381	ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr	400	Qy	741	MetLeuIleLeuLeuGlyGlnAlaGluAlaLeuLeuGluLysLeuValIleLeuHisAla	760
Db	1447	GTGGAAGCGCGCACCCACACAACTGAGGAGTGTCCGGCGGTATCCACCAATGGCAGTGG	1506	Db	2527	ATGCTCATCTTCTGGCGCAGGCTGGAAGCAGCAGCTAGAGAAGCTGCTTGTGCACGCT	2586
Qy	401	SerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTrp	420	Qy	761	AlaSerAlaAlaSerCysAsnGlyPheLeuTyrPheValIlePhePheValAlaAlaTrp	780
Db	1507	GGCATCTTCACCTCCGGCCCTAAGCAGACATTCAGCTCATTAACCAATGGCAGTGG	1566	Db	2587	GCGAGTGGGCTAGCTGCAATGGCCTCTGTATTTTGTCACTTTTTCGTGGCTGTCTGG	2646
Qy	421	HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer	440	Qy	781	TyrIleLysGlyArgValValProLeuAlaThrTyrSerLeuThrGlyLeuTrpSerPhe	800
Db	1567	CACATCAACCGCACTGCGCCTAAATGCAATGATCTTTGAACACCGGCTTTATGGCGCC	1626	Db	2647	TACATCAAGGTCGGCGGCTCCCTTGGCGCGCTACTCTCTCTACTGCGCTATGGCCCTTC	2706
Qy	441	LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg	460	Qy	801	SerLeuLeuLeuLeuAlaLeuProGlnGlnAlaTyrAlaTyrAspAlaSerValHisGly	820
Db	1627	TGTGTTCTACACCAAGAACTTCAACTGCTGGATGCGCCCGAACCGCTGTCGNCCTGCGC	1686	Db	2707	TGCTTACTCTCTAGTACTGCCCCAACAGGCTTATGCTTATGATGATCGTGCACGGA	2766
Qy	461	SerIleGluAlaPheArgValGlyTrpGlyAlaLeuGlnTyrGluAspAsnValThrAsn	480	Qy	821	GlnIleGlyAlaAlaLeuLeuValMetIleThrLeuPheThrLeuThrProGlyTyrLys	840
Db	1687	AACATCGAAGCTTCCCGATAGGATGGGGCACTTGGCAGTACGAGGATGATGTTACCNAT	1746	Db	2767	CAGTAGCGCGGCTTTGTTAATACTGATACCTCTTTCACACTTACCCCGGGGTATAG	2826
Qy	481	ProGluAspMetArgProTyrCysTrpHisTyrProProArgGlnCysGlyValValSer	500	Qy	841	ThrLeuLeuSerArgPheLeuTrpLeuCysTrpLeuLeuLeuThrLeuGlyGluAlaMet	860
Db	1747	CCAGAAGATATGAGACCATACTGCTGGCACTACCCACCAAAACAGTGTGGCATATTCGCC	1806	Db	2827	ACCCCTCTCAGTGGGTGTGTGGTGTGTGTCTCTCTGACCTTGGGGGAAGCCTTG	2886
Qy	501	AlaLysThrValCysGlyProValTyrCysPheThrProSerProValValValGlyThr	520	Qy	861	ValGlnGluTrpAlaProProMetGlnValLargGlyGlyArgAspGlyIleIleTrpAla	880
Db	1807	GCGGGGTCTGTGGCGGCCAGTGTACTGTTTCCACCCCGCCAGCCAGTAGTAGTGGGCACA	1866	Db	2887	ATCCAGAGTGGGTACACCCACCGCAGCGCGGTGCGCGCATGATCATATGGGCC	2946
Qy	521	ThrAspArgLeuGlyAlaProThrTyrThrTrpGlyGluAsnGluThrAspValPheLeu	540	Qy	881	ValAlaIlePheTyrProGlyValValPheAspIleThrIlystTrpLeuLeuAlaValLeu	900
				Db	2947	GCTACCATGTTCTGCCCGGGGTAGTGTGTTGACATAACCAAGTGGCTCTTGGCGGTGCTT	3006

QY 2981 LeuProGluAlaHargLeuLeuAapLeuSerSerTrpPheHrValGlyAlaGlyGlyGly 3000
 Db TTCCCGGAGGACGCTCTGGATTATTCAGTTGGTTACACCGTCGCGCGCGGGG 9306

QY 3001 AspIleTyHisSerValSerArgAlaAargProArgLeuLeuLeuPheGlyLeuLeuLeu 3020
 Db GACATTATACAGCGTGTGCGTGGCGGACCGCGCTTATTACTCTTGGCTACTCCTA 9366

QY 3021 LeuPheValGlyValGlyLeuPheLeuLeuProAlaAarg 3033
 Db CTTTCTGTAGCGTAGGCTCTTTCTACTCCCGCTCGG 9405

RESULT 12
 AF238482 9416 bp RNA linear VRL 26-MAR-2000
 DEFINITION Hepatitis C virus 2a polyprotein gene, complete cds.
 ACCESSION AF238482
 VERSION AF238482.1 GI:7329202
 KEYWORDS
 SOURCE
 ORGANISM
 Hepatitis C virus
 Hepatitis C virus
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.

REFERENCE
 1 (bases 1 to 9416)
 Itakura,J., Nagayama,K., Enomoto,N., Kurosaki,M., Watanabe,H. and Sato,C.
 Full length cDNA sequence of HCV genotype 2a, strain MD2a-2
 Unpublished

JOURNAL
 REFERENCE
 1 (bases 1 to 9416)
 Itakura,J., Nagayama,K., Enomoto,N., Kurosaki,M., Watanabe,H. and Sato,C.
 Direct Submission

JOURNAL
 Submitted (24-FEB-2000) Second Department of Internal Medicine, Tokyo Medical and Dental University, 1-5-45 Yushima Bunkyo-Ku, Tokyo 113-8519, Japan

FEATURES
 source
 1. .9416
 /organism="Hepatitis C virus"
 /mol_type="genomic RNA"
 /strain="MD2a-2"
 /db_xref="taxon:11103"
 /note="genotype: 2a"
 307. .9408
 /codon_start=1
 /product="polyprotein"
 /protein_id="AAF5941.1"
 /db_xref="GI:7329203"
 /translation="MSTNPKPQKTKRNTNRRPDVVFPGGQIVGGVYLLPRRGPRL
 GVRATKTSERSOPRRROPIDKRRSTGKSGKPGYWPVLYGNEGLGWAGWLLSPRG
 SRPSCPTDPRHSRNNGKVIDLTGCFADLMGYIPVCGAPLGGVARALAHGVRVLED
 GYNAVGNLPGCSFIFLLALLSCVTVPVSAVQVKNSTETVWTDNCSDNSITWQLOA
 AVLHPGCVPCERTGNTSCWIPVSNVAVQQGALTQGLRTHIDNVVMSATLCSALY
 VGDLCGVVLAALQFLTVSPRHHFVQECNSYPGAITHRWADMMNWSPTTMTVL
 AYAMRVEYIIDIISAGHWGFMFLAYFSMQAWAKVAVILLITAGVEARHTTTSVA
 GRTSGFAGIITSGPKONTQLINTGSHWNRNTALNCNDLNTGFMALPYTKNFNS
 GCPERLSKRNLEAFPIGWTLOVEDDVTNPRDMPCYCHYPPKOCGIFPAGSCGVP
 YCFPSPVVGTINKLGVITYWGENETDVFILNSTRPRGSGFWGCTWNSGTGFKTC
 GAPPCIRDFNASTDLCTCDPKRHPATYIKCGSPWLPRLCLLDVDPYRMLWHYPC
 TWNSYIFKRMVYGGVEHLRTACNFRSGRDNCLDRDSQTLPLHSTTEWAILPSC
 YDLPALSTGLLHLKIVDVQVMSQNGLLYFVIFVAAWYIKGRAVPLAAYSLTGLW
 MLILGQAALHLKIVLHAASACNGLLYFVIFVAAWYIKGRAVPLAAYSLTGLW
 PCCLLILVPOQAYDASVHGOVGAAILLITLTPGYKTLISRCILWCLYLTL
 GEALIQEVPPTQARGRGGIIWAATMFCGVVFDITKMLLAVLGEPAYLRLALTRVP
 YFVRAHALLRMKTVNLAGQVOMALLALRGWTGYIYDHLTPMSDWAAASGLRDLA
 VAVEPIIFSPMEKKVITANGAETAAACGDIKHGLPVSARLGREILLGPADGYISGWRL
 APITAYAGTQGLGAIIVVSMTRDKTEQAGEIQLVSTQSFGLTGISGLVWTVHG
 AGNKTLAGSRPVTOMYSAAEGDLVGMPPPGTKSLEPCTCGAVDLVLYTRNADVIPA
 RRGDKRGLISPRLLTKSGSGGVPVLCPRGHAVGIFRAAVCSRGVSGIDFIVET
 LDIVTSPTSDNSTPFAVPTQVGLHAPTSGSKSTKPVVAYAAQGYKVLVLPVS
 LATLGFAGVLSKAGHNPINRTGRTVTTGAALTYSTYKGLFLADGCGAGGAYDILIC
 ECHAVDSTILGVTDAQETAGVRLTVLATATPGSVTTPHNPTEEVALGEGEIP
 FYGRAIPLSHIKGGRHLIFCHSKKCDLAAALRGMGLNAVAYYRGLDVSVIPTQGDV
 VVPATDALMTGYTGDPDFIDCNVAVTVQVDFSLDPTFTITTTQTVQDVAVSRRQR

TGRLGIRYVYSTGERASGMFDSVVLCEYDAGAAWYELTFAETTVRLRAYENTPGL
 PYCQHLFEWAEVFTGLTHIDAFHSVOTKQSGENFAYLVAYQATVCARAKAPPSSDW
 MWKLTRLKPTLVGTELLYRSGVANEITLTHPTVTKYIATCQADLEWMTSVWLAG
 GYLAAYACLATGCVSIIGRLHINORAVAPDKVLYEAFDEMERCASRAALVEBQ
 RIAEMLSKIQGLLOQSKAQDIQPAVQASWPKVQFQWAKHWNFIQIYIAGLST
 VGENPAAVMFAFSAALTPSLSTTILLINILGMLASQIAPPAGATFVWGLVGA
 LFGILGKLVLDILAGYAGICALVAPKINSCEKPSMEDVNNLLPGIILSGALVGV
 ICAAILRRHVGEQAVOMNRLIAFASPKNSHAPTHYVTVESDASQVTLQGLSLTIT
 SLRLRHNIWITDCPIPCAGSLRDVMDWVCTILTDFKNWLSKLPKMPGLPFIISQ
 KGYKGMAGTGIMTRPCGANISGNVRLSGMRTGPKTCMTWQGTPIINCYTEQC
 VEKPAFNKVALRWASAEYVTHGYSYITGLTIDNLKVPQCLPSEFPFSSVNDVQ
 QIHRFAPTPKPFPRDEVSFCVGLNSFVGSQLPDDEPDDVLTLSMLTDPSHITAETA
 ARRLARGSPSEASASQISAPSLRATCTHGTVDVMDVMDASLFMGDVTIRSES
 KVVLDLSIDPMAEESDLEPSIPSEYMLSRNFPALPAWAPDYNPPLVPSWKRGLS
 QVATVAGCALPPKAPTPRRRTVGLSESTIGDALOOLAIKSPGQPPPSGSDGLS
 TGADAADSSRPPDELALESETGSSMPPELGEPCDPLEPEQVPPQPPQGEVAP
 GDSGSWSTCSEENVSVFCSISYWTGGLITPCIPEKELKPNLUSNLLIDYDNKYI
 CTTSKASLRAKKVTDFDRMOVLDAVYDSDIKLAASKVSARLLLEBACQITPPHS
 ARSKYGFCAKEVRSISGRAVNHKSVMKOLLEDSDPTPTTMAKNEVFCVDPKGGK
 KAARLIVPDLGVCEKMAVDYTKLPQAVMGASYGFGYPAQVFEVLLKAWAGK
 DMWGSYDTRCFDSTVTERDIRTEESIVRPSCLPREAHTAIHSLTERLYVGGPMFNSK
 GQACGYRRCRCSVLTSMGNTITCYVKAALACKAGIIVAPTMVLCDDLDVLYSESQ
 TEEDERNURKAFTEAMTRYSAPEDPPEIDLELITSSSNVSHVHGGQRRRIYLR
 DPTPIARAWEETVRHSVNSWLGNIIOYAPTVMVWMLMTHFFSLMAQDTLDQNLN
 FEMYGSVYSVPLDLPAIIRLHGLDAFSLHTYSPHELTRVAAARKLGPAPLRANKS
 RARAVRASILSHGGRAAVCGRYLFNVAVKTKLITPLPEARHLDLSSWFTVGAGGCHI
 YHVSRSRPRLLGLLLLFVCGVGLFLPAR"

BASE COUNT 1930 a 2790 c 2652 g 2042 t 2 others
 ORIGIN

Alignment Scores:
 Pred. No.: 0 Length: 9416
 Score: 15355.00 Matches: 2857
 Percent Similarity: 96.93% Conservative: 83
 Best Local Similarity: 94.20% Mismatches: 93
 Query Match: 95.25% Indels: 0
 DB: 14 Gaps: 0

US-09-980-559-2 (1-3033) x AF238482 (1-9416)

QY 1 MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGln 20
 Db 307 ATGAGCACAAATCTTAAACCTCAAGAGAAAACCAAGAAATACTAATACCGTCCCCACAG 366

QY 21 AspValLysPheProGlyGlyGlnIleValGlyValTyrlleuLeuProAarg 40
 Db 367 GAGCTCAAGTTTCGCGCGCGCGCGCAGATCGTTGGCGGAGTATACTTGTTCGCGCGAGG 426

QY 41 GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly 60
 Db 427 GACCCCAAGTTGGGTGTGCGCGCGCACAGAGAGACTTCGAGCGGCTCCAGCACACGTGGG 486

QY 61 ArgArgGlnProIleProLysAspArgArgSerThrGlyLysSerTrpGlyLysProGly 80
 Db 487 AGCGCGCAGCCCATCCCAAGATCGCGCTCCACTGGCAAGTCTCTGGGGAACCAACAGGA 546

QY 81 TyrProTrpProLeuTyrlGlyAsnGluGlyLeuGlyTrpAlaGlyTrpLeuLeuSerPro 100
 Db 547 TACCCCTGGCCCTATATGGGAATGAGGAGACTCGCGTGGCGGAGGATGGCTCTGTCCCC 606

QY 101 ArgGlySerArgProSerTrpGlyProAsnAspProArgHisArgSerArgAsnValGly 120
 Db 607 CAGAGTTCCCGTCCCTCTTGGGGCCCCCCTGACCCCCCGGCATAGGTTCGGCAACGTGGGT 666

QY 121 LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrlleProValVal 140
 Db 667 AAGGTATCATGATACCTAACGTGTGGCTTTTGGCGGCTCATGGGGTACATCCCGCTCGTA 726

QY 141 GlyAlaProLeuGlyGlyValalalargAlaLeuAlaHisGlyValArgValLeuGluAsp 160
 Db 727 GCGGCCCACTTGGTGGCGTTCACAGAGCTCTCGCGCACGCGGTGAGAGTCTCTGGAGGAC 786

QY 161 GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla 180

7027 GTGACTCGGATAGATCTCAGTCAAAAGTGGTCTTCTCGGACTCTCTCGACCAATGGCT 7086
 2261 GluGluArgSerAspLeuGluProSerIleProSerGluTyrMetLeuProLysLysArg 2280
 7087 GAAGAAAGAGGACCTCGAGCCCTCAATACCATCGAGTATATGCTCCCCAGGAACAGG 7146
 2281 PheProProAlaLeuProAlaTrpAlaArgProAspTyrAsnProProLeuValGluSer 2300
 7147 TTCCCGCCAGCCTTACCGGCTGGGCGCGCTGATTAACCCACCCGCTTGTGGNATCG 7206
 2301 TrpLysArgProAspTyrGlnProAlaThrValAlaGlyCysAlaLeuProProArg 2320
 7207 TGAAGAGGCGCAGATTACCAACCGCCACTGTTGCGGCTGTCTCTCCCGCCCTAAG 7266
 2321 LysThrProThrProProArgArgArgThrValGlyLeuSerGluAspSerIle 2340
 7267 AAAACCCCGACGCTCCCGAAGGCGCGGACAGTGGTCTGGATGAGACCACTA 7326
 2341 GlyAspAlaLeuGlnGlnLeuAlaIleLysSerPheGlyGlnProProProSerGlyAsp 2360
 7327 GGAGATGCCCTCAGCAGCTGGGCATCAAGACCTTCGGCCAGATCCCGCCACGCGCAT 7386
 2361 SerGlyLeuSerThrGlyAlaGlyAlaAspSerGlySerGlnThrProProAspGlu 2380
 7387 TCAGGCTTTCCACGGGGCGGCGCGCGACTCCGGCGCGGCGGCGTCCCTGACGAG 7446
 2381 LeuAlaLeuSerGluThrGlySerIleSerSerMetProProLeuGluGlyLeuGly 2400
 7447 TTGGCTCTTTCCGAGACAGGTTCCATTTCTCCATGCCCGCCCTTGAGGGGGAGCCTGGG 7506
 2401 AspProAspLeuGluProGluGlnValGluProGlnProProGlnGlyValAla 2420
 7507 GATCCAGACTAGAGCCCGAGAGGTAGACTTCAACTCCCGCCCGAGGGGGAGCTA 7566
 2421 AlaProGlySerAspSerGlySerTrpSerThrCysSerGluGluAspSerValVal 2440
 7567 GCTCCCGGCTCGGACTCGGGGCTCTGGTCTGCTGCTCCGAGGAGGATGACTCCGTCGTG 7626
 2441 CysCysSerMetSerTyrSerTrpThrGlyAlaLeuIleThrProCysSerProGluGlu 2460
 7627 TGCTGCTCCATGCTACTCTGACCGGGCTCTAATACTCTTGTAGCCCGCGAAGAA 7686
 2461 GluLysLeuProIleAsnProLeuSerAsnSerLeuLeuArgTyrHisAsnLysValTyr 2480
 7687 GAGAAGTTCAATTAACCTCTGAGCACTCGTGTGGATATCAACACAGGTATAC 7746
 2481 CysThrThrLysSerAlaSerLeuArgAlaLysLysValThrPheAspArgMetGln 2500
 7747 TGTACTACATCAAGAGTGGCTCACTGAGGGCTAAAAAGGTAACTTTGACAGGATGCAA 7806
 2501 ValLeuAspSerTyrTyrAspSerValLeuLysAspIleLysLeuAlaSerLysVal 2520
 7807 GTGCTCGACGCCATTAAGACTAGTCTTAAGAGACATCAAGCTAGCGGCTCTAAGGTC 7866
 2521 ThrAlaArgLeuLeuThrMetGluGluAlaCysGlnLeuThrProProHisSerAlaArg 2540
 7867 AGTCAAGGCTCTCACTTAGAGAGGCGCTGACGTTGACCCACCCCATCTGCAAGA 7926
 2541 SerLysTyrGlyPheGlyAlaLysGluValArgSerLeuSerGlyArgAlaValAsnHis 2560
 7927 TCCAAATATGATTTGGGCTAAAGAGGTCCGAGCTTGTCCGGGAGAGCGCTCAACAC 7986
 2561 IleLysSerValTrpLysAspLeuLeuGluAspSerGluThrProIleProThrIle 2580
 7987 ATCAAGTCCGTTGGAGGACCTCTCGAAGACTCACAAACACCAATTCCTTACGACCATC 8046
 2581 MetAlaLysAsnGluValPheCysValAspProThrLysGlyLysLysAlaAlaArg 2600
 8047 ATGCCCCAAATAGAGTGTCTGCTGCACCCACACAGGGGGGTAGAAAGCAGCTCGC 8106
 2601 LeuIleValTyrProAspLeuGlyValArgValCysGluLysMetAlaLeuTyrAspIle 2620
 8107 CTTATCGTTTACCCTGACCTCGGCTGAGGCTATGCGAGAAATGGGCCCTTTATGATATC 8166

2621 ThrGlnLysLeuProGlnAlaValMetGlyAlaSerTyrGlyPheGlnTyrSerProAla 2640
 8167 ACACAAAAGCTTCTCAGCGGTGATGGGGCTTCTTATGGCTTCCAGTACTCCCTGCA 8226
 2641 GlnArgValGluPheLeuLeuLysAlaTrpAlaGluLysLysAspProMetGlyPheSer 2660
 8227 CAACGGTGGAGTTTCTCTGAAGCATGGGGGAAAGAACACCCATTTGGTTTTCG 8286
 2661 TyrAspThrArgCysPheAspSerThrValThrGluArgAspIleArgThrGluGluSer 2680
 8287 TATGATACCCGATGCTTTCAGCTCAACCGTCACTGAGAGAGCATCAGAACTGAGGATCC 8346
 2681 IleTyrArgAlaCysSerLeuProGluGluAlaHisThrAlaIleHisSerLeuThrGlu 2700
 8347 ATATACCGGCTCTCTCCCTGCGGAGGAGCTCACTGCTCATACATTCGCTGACTGAG 8406
 2701 ArgLeuTyrValGlyGlyProMetPheAsnSerLysGlyGlnThrCysGlyTyrArgArg 2720
 8407 AGACTTTACGTTGGGGGGCTTATGCTCAACAGCAAGGCCAATCTTCGCGGTACAGGCT 8466
 2721 CysArgAlaSerGlyValLeuThrThrSerMetGlyAsnThrIleThrCysTyrValLys 2740
 8467 TGCGCGCGCAGCGGAGTGTCTACCACTAGCATGGGGAACACCATCACATGCTATGTGAA 8526
 2741 AlaLeuAlaAlaCysLysAlaAlaGlyIleIleAlaProThrMetLeuValCysGlyAsp 2760
 8527 GCCCTAGCAGCTTGTAAAGCTGCGGGATAGTTGGGCCCAATGCTGCTATGCGCGAC 8586
 2761 AspLeuValValIleSerGluSerGlnGlyThrGluGluAspGluArgAsnLeuArgAla 2780
 8587 GACTTGGTGTCTCATCTCAGAAAGCCAGGGGACTGAGGAGGAGCGAGCAACCTGAGAGC 8646
 2781 PheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProArgProGluTyr 2800
 8647 TTCACGAGGCTATGACCGAGGTATCTGCCCTCTCTGGTGACCTCCAGACCGGATAT 8706
 2801 AspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaLeuGlyProGlnGly 2820
 8707 GACTGGAGCTGATAACATCTTGTCTCTCAATGTCTCGTGGCGTGAACCCACCTGGGC 8766
 2821 ArgArgTyrTyrLeuThrArgAspProThrThrProIleAlaArgAlaAlaTrpGlu 2840
 8767 CGCGCGAGACTACTACCTGACGAGACCCACCATCCAAATTCGCGGGCTGCTGGAA 8826
 2841 ThrValArgHisSerProValAsnSerTrpLeuGlyAsnIleIleGlnTyrAlaProThr 2860
 8827 ACAGTTAGACACTCCCTCTCAATTCATGGCTGGGAAACATCATCCAGTACGCCCAACA 8886
 2861 IleTrpAlaArgMetValLeuMetThrHisPheSerIleLeuMetAlaGlnAspThr 2880
 8887 ATATGGCTCGCATGCTCTGATGACACACTTCTCTCCATTCATGGCTCAAGACACC 8946
 2881 LeuAspGlnAsnLeuAsnPheGluMetTyrGlyAlaValTyrSerValSerProLeuAsp 2900
 8947 CTGGACAGAACTCTCACTTTGAGATGTCGGGTACTCCGTGACTCCGTGAGTCTCTTGGAC 9006
 2901 LeuProAlaIleIleGluArgLeuHisGlyLeuAspAlaPheSerLeuHisThrTyrThr 2920
 9007 CTCCAGCCATAATTGAAGGCTACAGGCTTTCAGCTTCTCTCTGACACATATCTCT 9066
 2921 ProHisGluLeuThrArgValAlaSerAlaLeuArgLysLeuGlyAlaProProLeuArg 2940
 9067 CCCACGAACTGACCGGGTGGTTCACCCCTCAGAAATCTGGGGCGCCACCCCTCAGA 9126
 2941 AlaTrpLysSerArgAlaArgAlaValArgAlaSerLeuIleSerArgGlyGlyArgAla 2960
 9127 GCGTGAAGAGTCCGGCGCGTGCAGTGTAGGCGCTCCCTCATCTCNCGTGGAGGAGAGCG 9186
 2961 AlaValCysGlyArgTyrLeuPheAsnTrpAlaValLysThrLysLeuThrPro 2980
 9187 GCCGTTTGGGCTCGGTATCTCTTCACTGGCGGTGAAGCAACCAAGCTCAAACTCACCTCA 9246

QY 1521 LeuCysGluCysTyrAspAlaGlyAlaAlaTyrTyrGluLeuThrProSerGluThrThr 1540
DB 4867 CTCTGTGAGTGTACGACGAGGAGTGTGATGTATGAGCTTACACGAGGAGACAC 4926
QY 1541 ValArgLeuArgAlaTyrPheAsnThrProGlyLeuProValCysGlnAspHisLeuGlu 1560
DB 4927 GTGAGGCTTAGAGCTTATTTTAAACACACCTGGCTTGGCTGTGGCAAGACCATTTGAG 4986
QY 1561 PheTyrGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThr 1580
DB 4987 TTTTGGGAGGAGCTTTTACCGGCTTACACATAGATGCCACTTCTTTCCTCCAAACA 5046
QY 1581 LysGlnSerGlyGluAsnPheAlaTyrLeuThrAlaTyrGlnAlaThrValCysAlaArg 1600
DB 5047 AAGCAAGCGGAGAGAGCTTTTCCATACCTGGTACCTTATCAAGCTACAGTATGCGCTAGA 5106
QY 1601 AlaLysAlaProProSerThrAspValMetTyrPysCysLeuThrArgLeuLysPro 1620
DB 5107 GCCAAAGCCCCCGCTCTGGGACATTTATGTGGAAGTGTAACTCGACTCAAGCCCC 5166
QY 1621 ThrLeuValGlyProThrProLeuLeuTyrArgLeuGlySerValThrAsnGluValThr 1640
DB 5167 ACCTGTGGGCGCTTACACCCCTTCTGTACCGTTTGGCTCTGTACCAAGAGATCACC 5226
QY 1641 LeuThrHisProValThrLysTyrIleAlaThrCysMetGlnAlaAspLeuGluValMet 1660
DB 5227 CTCACACATCTGTGAGGAATACATCGCCACATGCATGCAAGCTGACCTTGAGGTCATG 5286
QY 1661 ThrSerThrTyrValLeuAlaGlyGlyValLeuAlaValAlaValAlaTyrCysLeuAla 1680
DB 5287 ACCAGACGCTGGGTCTTAGCTGGGGAGTCTTGGCAGCGCTCGCCGCTATTGCTGGCG 5346
QY 1681 ThrGlyCysValCysIleIleGlyArgLeuHisIleAsnGlnArgAlaValValAlaPro 1700
DB 5347 ACCGGTGTGTTTCCATCATCGCGCTTGCATATCAATCAGGAGCCATCGTGTGACCG 5406
QY 1701 AspLysGluValLeuTyrGluAlaPheAspGluMetGluGluCysAlaSerArgAlaAla 1720
DB 5407 GACAAGGAGTCTCTTTATGAGGCTTTTGTAGATGAGAGGAATGTGCTCTAAAGCGGCT 5466
QY 1721 LeuIleGluGlyGlnArgIleAlaGluMetLeuLysSerLysIleGlnGlyLeuLeu 1740
DB 5467 CTCCTAGAAGAGGGCAGCGAATAGCCGAGATTTGAAAGTCCAAGATCCAAGGCTTACTG 5526
QY 1741 GlnGlnAlaSerLysGlnAlaGlnAspIleGlnProThrValGlnAlaSerTyrProLys 1760
DB 5527 CAGCAAGCTCTTAAACAGGCCAGGACATACAAACCCGCTGTGAGGCTTCATGCGCCAAAG 5586
QY 1761 ValGluGlnPheThrAlaLysHisMetTyrAsnPheIleSerGlyIleGlnTyrLeuAla 1780
DB 5587 GTGGAACAATTCCTGGGCCAAACACATGTGGAACCTTCATTAGCGGCATCCATACCTAGCA 5646
QY 1781 GlyLeuSerThrLeuProGlyAsnProAlaValAlaSerMetAlaPheSerAlaAla 1800
DB 5647 AGACTGTCAACACTGCGCAGGAACCCCGCTGTGGCTTCATGATGCGCTTCAGTGGCGCC 5706
QY 1801 LeuThrSerProLeuSerThrThrIleLeuLeuAsnIleLeuGlyGlyTyrLeu 1820
DB 5707 CTCACAGCCCGGTGTCAACTAGCACCATCTCTCTCAACATTTTGGGAAGGTGGCTG 5766
QY 1821 AlaSerGlnIleAlaProProAlaGlyAlaThrGlyPheValValSerGlyLeuValGly 1840
DB 5767 GCGTCCCAATTTGCTCCACCGCGGGGGCCACCGGCTTGTGTGATGCTGTGGTGGGG 5826
QY 1841 AlaAlaValGlySerIleGlyLeuGlyLysValLeuValAspIleLeuAlaGlyTyrGly 1860
DB 5827 GCTGCTGTGGCAATATAGTCTGGGTAAGGTGCTGTGGACATTTGCGAGGATATGCG 5886
QY 1861 AlaGlyIleSerGlyAlaLeuValAlaPheLysIleMetSerGlyLysProSerMet 1880
DB 5887 GCGGGCATTTTGGGGGCGCTTGTGCGCATTCAGATCATGTGTGGCGAGAGCCCTCCATG 5946
QY 1881 GluAspValValAsnLeuLeuProGlyIleLeuSerProGlyAlaLeuValGlyVal 1900

DB 5947 GAAGATGTCAATCAACCTGTGCTGGGATCTCTGCCCGGGTCCCTGTGTGGTGGGGTC 6006
QY 1901 IleCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTyrMet 1920
DB 6007 ATCTGGCGCGCATTTCTGGCGCCACGTGGGACCGGGGGAAGCGCGGTCCAAATGGATG 6066
QY 1921 AsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaProThrHisTyrValThr 1940
DB 6067 AACAGGCTCATCGCTTCGCTTCCAGAGGAACACCGTCCGCCCTACTCCTACGTAGTCAGC 6126
QY 1941 GluSerAspAlaSerGlnArgValThrGlnLeuLeuGlySerLeuThrIleThrSerLeu 1960
DB 6127 GAGTCAGATGCGTGCAGCGTGTGACCCAGCTGCTGGCTCTCTTACAATAACTAGCCTA 6186
QY 1961 LeuArgArgLeuHisAsnTyrIleThrGluAspCysProIleProCysGlyGlySerTyr 1980
DB 6187 CTCAGGAGACTCCACAATTTGGATTTACTGAGGACTGCCCATGCCCGCGCTCGTGG 6246
QY 1981 LeuArgAspValTyrAspTyrValCysThrIleLeuThrAspPheLysAsnTyrLeuThr 2000
DB 6247 CTTGGGATGTGGGACTGGGTTGCACCATCTGACAGATTTTAAAACTGGCTGACT 6306
QY 2001 SerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTyrLys 2020
DB 6307 TCCAAAGTTCTCCAAAGATGCTGGCCTCCCTTTATCTCTTCCCAAAAGGGGTACAAG 6366
QY 2021 GlyValTyrAlaGlyThrGlyIleMetThrThrArgCysProCysGlyAlaAsnIleSer 2040
DB 6367 GGTGTGTGGCGCGCATCTGATATCATGACACACGATGTCCCTCGCGCGCCCAACTTCT 6426
QY 2041 GlyAsnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTyr 2060
DB 6427 GGCATGTCCGCTGGGCTCAATGAGATCACAGGSCCTAAACCTGCATGATGAACATTGG 6486
QY 2061 GlnGlyThrPheProIleAsnCysTyrThrGluGlyGlnCysValProLysProAlaPro 2080
DB 6487 CAGGGAGACCTTCTTATCACTGTTTACACTGAGGGCCAGTGGTGGCCGAAACCGCGCG 6546
QY 2081 AsnPheLysValAlaIleTyrArgValAlaAlaAsnGluTyrAlaGluValThrGlnHis 2100
DB 6547 AACTTCAAGACCGCATCTCGAGGGTGGCGGCTCAGAGTACGCGAGGTGACGCGACAC 6506
QY 2101 GlySerTyrHisTyrIleThrGlyLeuThrAspAsnLeuLysValProCysGlnLeu 2120
DB 6607 GGGCAATCTCTATGTGACAGATTGACCATGCAACTTGAAAGTTCCCTGCCAACTA 6666
QY 2121 ProSerProGluPhePheSerTyrValAspGlyValGlnIleHisArgPheAlaProThr 2140
DB 6667 CCTTCTCCAGAGTCTCTTCTTCTGGTGGAGCGAGTGCAAAATCCATAGGTTTGGCCCA 6726
QY 2141 ProLysProPhePheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal 2160
DB 6727 CCAAGACCGTCTTCTCGGATGAGGTCTGTTTATTTATTTGGGCTCAATTCGTTTGTGTA 6786
QY 2161 GlySerGlnLeuProCysAspProGluProAspThrAspValIleuMetSerMetLeuThr 2180
DB 6787 GGGTCTCAGCTCCCTGTGATCTTCTGAGCCCGATACAGATGTGTGATGTCCATGTTTAA 6846
QY 2181 AspProSerHisIleThrAlaGluThrAlaAlaArgArgLeuAlaArgGlySerProPro 2200
DB 6847 GAACCATCCCATATTACGGCGAGGCTGCGCGCGCTTAGCAGCGGGGTCAACCCCG 6906
QY 2201 SerGluAlaSerSerAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThr 2220
DB 6907 TCTGAGGCGAGCTCTTCAGCGAGCGAGCTATCAGCACCATCTGCTGCGAGCCACTGCACC 6966
QY 2221 ThrHisGlyLysAlaTyrAspValAspMetValAspAlaAsnLeuPheMetGlyGlyAsp 2240
DB 6967 ACCCAGCGCAAGGCTTATGATGTGACATGTGGATGCGCACTGTTTCAATGGGGCGCAT 7026
QY 2241 ValThrArgIleGluSerGlySerLysValValLeuAspSerLeuAspProMetVal 2260

1161	ProfLeuSerThrLeuLysGlySerSerGlyClyProValLeuLysProArgGlyHisAla	1180
3787	CCCTTTTCGACCTTGAAGGGGTCTCGGGGGACCGGTGTCTTGCCCTTAGGGGCCACGCT	3846
1181	ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle	1200
3847	GTCCGGATCTTCGGGCAGCTGTGTCTCTCGGGCGTGGCTAAGTCCATAGATTTCATC	3906
1201	ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro	1220
3907	CCCGTTGAGACACTTGACATCGTCACCGGGTCTCCACACTTTTAGTGACAACAGCACCA	3966
1221	ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLys	1240
3967	CCGGCTGTGCCCCAGACCTATCAGGTCCGGTACTTACATGCCCCCAACTGGCGAGTGGGAAG	4026
1241	SerThrLysValProValAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnPro	1260
4027	AGCACCAAGTCCCTGTGCGCTAGCGCGCTCAGGGGTATAAGGTACTAATTCTTAATCCC	4086
1261	SerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerLysAlaHisGlyIleAsnPro	1280
4087	TCGGTGGGTGCCACCTCTGGGGTTTGGGGCGTACTTGTCCAAGGCATATGGCATCAATCCC	4146
1281	AsnIleArgThrGlyValArgThrValThrThrGlyAlaProIleThrTyrSerThrTyr	1300
4147	AACATCAGAGCTGGGTCCAGAACTGTGCACACCGGGGCTCCATCACAATCCACATAT	4206
1301	GlyLysPheLeuAlaAspGlyGlyCysAlaGlyGlyAlaTyrAspIleIleIleCysAsp	1320
4207	GGTAAATCTTCGCTGATGGGGCTGCGCAGCGCGGCTATGACATCATCATATGCGAT	4266
1321	GluCysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGlnAla	1340
4267	GAATGCCACGCGTTGGATCTACCAACCAATCTTCGGTATCGGAACAGTCTCTGATCAAGCA	4326
1341	GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProProGlySerValThr	1360
4327	GAGACGGCGGGGTCCAGACTAACTGTGTGTGGTACGGCTACGCTCCCGGGTCAGTGGAG	4386
1361	ThrProHisProAsnIleGluAluAlaLeuGlyGlnGluGlyValIleProPheTyr	1380
4387	ACCCGCCCATCCCAATATAGAGAGGTAGCCCTCGGGCAGAGGGTGAGATCCCTCTCTAT	4446
1381	GlyArgAlaIleProLeuSerTyrIleLysGlyIleArgHisLeuIlePheCysHisSer	1400
4447	GGGAGGGCGATTCCTCTGCTTACATCAAGGGAGGGAGGCACCTTGATCTTCTGCCACTCA	4506
1401	LysLysLysCysAspGluLeuAlaAlaLeuArgGlyMetGlyLeuAsnSerValAla	1420
4507	AAAGAAAAGTGTGACGAGCTCCGGCGCGCCCTTCGGGGCATCGGGCTTGAACGTGTGGCA	4566
1421	TyrTyrArgGlyLeuAspValSerValIleProThrGlnGlyAspValValValAla	1440
4567	TACTACAGAGGGTTGACGTCTCCGTAATACCACATCAGGAGAGTGTAGTGTGTCTGCC	4626
1441	ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnVal	1460
4627	ACCGACGCCCTCATGACGGGGTTCACTGGAGACTTTTGACACCGGTGATCGACTGCCAACGTA	4686
1461	AlaValThrGlnValValAspPheSerLeuAspProThrPheThrIleThrThrGlnIle	1480
4687	GGGGTCACTCAAGTCGTAGACTTCAGCTTGGACCCGCCCTTACCATAACCAACACAGACT	4746
1481	ValProGlnAspAlaValSerArgSerGlnArgArgGlyArgThrGlyArgGlyArgLeu	1500
4747	GTACCTCAAGACGCTGTCTCAGTAGCCAGCGCGGGGGCGCAACGGGTAGAGGAAGACTG	4806
1501	GlyIleTyrArgTyrValSerThrGlyLysArgAlaSerGlyMetPheAspSerValVal	1520
4807	GGCAGCTTATAGGTACGTTTCCATCGTGTGAACAGCTTTCAGAAATGTTTGAAGTGTAGT	4866

QY 61 ArgArgGlnProIleProIleSerArgSerThrGlySerThrGlySerThrGlySerProGly 80
 Db 487 AGCGCGCAGCCATCCCAAGATCGCGCTCCACTGGCAAGTCTCTGGGGAACACAGGA 546
 QY 81 TyrProTrpProLeuTyrGlyAsnGluGlyLeuGlyTyrAlaGlyTyrLeuLeuSerPro 100
 Db 547 TACCCCTGGCCCTATATGGGAATGAGGACTCGGCTGGGCGAGGATGGCTCTGTCTCCCG 606
 QY 101 ArgGlySerArgProSerTrpGlyProAsnAspProArgHisArgSerArgSerValGly 120
 Db 607 CGAGGTTCCCGTCCCTCTTGGGGCCCTACTGACCCCGGCTAGGTTCGGCGAACGTGGGT 666
 QY 121 LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProValVal 140
 Db 667 AGGTTCATCGATACCTACGTGGGCTTTCGGGCTCATGGGCTACATCCCGTCTGTA 726
 QY 141 GlyAlaProLeuGlyGlyValAlaAlaAlaLeuAlaHisGlyValArgValLeuGluAsp 160
 Db 727 GCGCGCCCGCTTGGTGGCGTGGCAGAACTCTCGCGCAGCGGCTGAGAGTCTCTGGAGGAC 786
 QY 161 GlyValAspPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla 180
 Db 787 GGGGTTAAATATGCAACAGGAACTTACCTGGTGTCTCTTCTATCTTCTTGTGGCC 846
 QY 181 LeuLeuSerCysIleThrThrProValSerAlaAlaGluValLysAsnIleSerThrGly 200
 Db 847 CTACTGTCTGCATCACCATTCCGGTCTCTGCTATCGAAGTGAAGAACACGAGTACGCT 906
 QY 201 TyrMetValThrAsnAspCysThrAsnAspSerIleThrTyrGlnLeuGlnAlaVal 220
 Db 907 TATATGGTGAATAATGACTGTTCCATGGCAGCANTCACCTGGCACTCCAGGCGCGGTC 966
 QY 221 LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTrpIle 240
 Db 967 CTCACGTCCTCCCGGGTGGCTCGCGTGGAGAAAGTGGGAAATACATCACGGTGTGGATA 1026
 QY 241 ProValSerProAsnValAlaValAlaGlnArgProGlyAlaLeuThrGlnGlyLeuArgThr 260
 Db 1027 CCGGTTCTACCAACAGTAGCGGTGACGGCTGGCGCTCTACGCGAGGCGCTTGGCGAGC 1086
 QY 261 HisIleAspMetValValMetSerAlaThrLeuCysSerAlaLeuTyrValGlyAspLeu 280
 Db 1087 CACATCGCATGGTGTGTATGTCGCGCACGCTCTGCTCTCTCTACGCTGGGGGACCTC 1146
 QY 281 CysGlyClyValMetLeuAlaAlaGlnMetPheIleValSerProGlnHisIleThrPhe 300
 Db 1147 TCGCGGGGGTGTATGTCGCGACCCCATATGTTTATTGTCACGCGCAACACCACTGGTTT 1206
 QY 301 ValGlnAspCysAsnCysSerIleTyrProGlyThrIleThrGlyHisArgMetAlaTrp 320
 Db 1207 GTGCGAGATTGCNATTTGCTTCCATCTACCTTGTGTACCATCTCTGGACACCGCATGG 1266
 QY 321 AspMetMetMetAsnTrpSerProThrAlaThrMetIleLeuAlaTyrAlaMetArgVal 340
 Db 1267 GACATGATGATGAATGGTTCGCGCACCGCACCATCATCTGGCGTACGCGATGCGGCTC 1326
 QY 341 ProGluValIleIleAspIleIleSerGlyAlaHisTrpGlyValMetPheGlyLeuAla 360
 Db 1327 CCGGAGGTTCATCATAGACATCATTTGGTGGGGTTCATTGGGGGCTCATGTTTCGGCTTAGCC 1386
 QY 361 TyrPheSerMetGlnGlyAlaTrpAlaLysValValIleLeuLeuLeuAlaAlaGly 380
 Db 1387 TACTTCTCTATGAGGAGGAGCTGGGCAAAAGTCTTGTCTCTCTCTCTGCTGACCGTGGG 1446
 QY 381 ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr 400
 Db 1447 GTGAGCGCAATACCTACTCCGTTCGGGGTTCCTACTGCGCATCTCACCCAGACCTTTGCC 1506
 QY 401 SerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTrp 420
 Db 1507 GGCCTTTCACCTGGGGCTCTCAGCAGAAATATCCAGCTCATCAACCAATGGCGAGTTGG 1566
 QY 421 HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer 440

Db 1567 CACATCAACCGTACTGCCCTGAATTCGAATCCCTCAACACCGGCTTCTTCGGCTCC 1626
 QY 441 LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg 460
 Db 1627 CTGTTCTACACCCAGCGGTTCAACTCGTCAGGATGTCGCGAAGCGCTTCGCGCTCCGCG 1686
 QY 461 SerIleGluAlaPheArgValGlyTyrGlyAlaLeuGlnTyrGluAspAsnValThrAsn 480
 Db 1687 AACATCGAGGCTTTCCGATAGATGGGGCGTTTTCGAATACGAGGAATATGTCAACAT 1746
 QY 481 ProGluAspMetArgProTyrCysTrpHisTyrProProArgGlnCysGlyValValSer 500
 Db 1747 CCAGAAGATATGAGACCATATTTGTCGCACTACCCACCAAAACAGTCGCGCATAGTCCCC 1806
 QY 501 AlaLysThrValCysGlyProValTyrCysPheThrProSerProValValValGlyThr 520
 Db 1807 GCGAGGCTCTGTATGTGGCGGCTGTACTGTTTCACTCCAGCCAGTGTGTGGGACG 1866
 QY 521 ThrAspArgLeuGlyAlaProThrTyrThrTyrGlyGluAsnGluThrAspValPheLeu 540
 Db 1867 ACCGATAGCTTGGAGTGCCTACTTATCATGGGAGAGAAATGAGACAGATGTCTTCTTA 1926
 QY 541 LeuAsnSerThrArgProProLeuGlySerTrpPheGlyCysThrTrpMetAsnSerSer 560
 Db 1927 TTGAATAGCACCCGCGCCACCGCGGGGCGCATGGTTCGGCTGCACGTGGATGAATCCACT 1986
 QY 561 GlyTyrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer 580
 Db 1987 GGTTCACCAAGACTTGTGGCGCACCCCTTCGCGCATCAGAGCTGACTTCAACGCGCAGT 2046
 QY 581 ThrAspLeuLeuCysProThrAspCysPheArgLysHisProAspThrThrTyrLeuLys 600
 Db 2047 ACGGACCTGTATGCCCCACGACTGTTTTAGAGACATCTTGAGGCCACTTACACCAA 2106
 QY 601 CysGlySerGlyProTrpLeuThrProArgCysLeuIleAspTyrProTyrArgLeuTrp 620
 Db 2107 TGTGGCTCTGGGCTTGGCTCAGCGCAAGATCATGGTCGACTACCTTACAGGCTCGGC 2166
 QY 621 HisTyrProCysThrValAsnTyrThrIlePheLysIleArgMetTyrValGlyGlyVal 640
 Db 2167 ATTCTCCCTGCACAGTCAACTATACCATCTTCAAGATAGGATGTATGTGGGGGGT 2226
 QY 641 GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp 660
 Db 2227 GAACACAGCTTACGCGCGGTGTAATCTCTCTGTTGGGGATCTGTGCAACTTGGAGGAC 2286
 QY 661 ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGluTrpAlaIleLeuPro 680
 Db 2287 AGAGACAGAAGTCAACTGTCTCTTTGTTGCACTCCACCGAATGGGCCATTTTGCCT 2346
 QY 681 CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuHisLeuHisGlnAsnIle 700
 Db 2347 TGTACTTCTCAGACTTGGCGCTTGTGCACTGGTCTTCTCCACCTCCACCAAAACATC 2406
 QY 701 ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrLysTyrIleValArgTrp 720
 Db 2407 GTGGAGGTGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2466
 QY 721 GluTrpValIleLeuLeuPheLeuLeuAlaAspAlaArgValCysAlaCysLeuTrp 740
 Db 2467 GAGTGGGTGATCTCTTATCTCTGCTTGGCGGAGCGCCAGGCTCTGCGCATGCGTGG 2526
 QY 741 MetLeuIleLeuLeuGlyGlnAlaGluAlaAlaLeuGluLysLeuValIleLeuHisAla 760
 Db 2527 ATGCTCATCTTCTGGCGGAGCGCCAGCAGCAGTACAGAGCTGTGCTGTCTTTCACGCT 2586
 QY 761 AlaSerAlaAlaSerCysAsnGlyPheLeuTyrPheValIlePhePheValAlaAlaTrp 780
 Db 2587 GCGAGCGCAGTGTGCAATGGCTTCTTATATTTGTCTTCTTCTTCTTCTTCTTCTTCTT 2646
 QY 781 TyrIleLysGlyArgValValProLeuAlaThrTyrSerLeuThrGlyLeuTrpSerPhe 800

6761	CGAAGCGCTTTTCGGGATAGGCTCGTTTGGCTCGGGCTTAACATTTGTCGTC	6820	7841	GTGCTCGAGCGCTTTACGACTCAGTCTTAAAGGACATCAAGCTAGCGGCTCAAGGTC	7900
2161	GlySerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThr	2180	2521	ThrAlaArgLeuLeuThrMetGluGluAlaCysGlnLeuThrProProHisSerAlaArg	2540
6821	GGGTCTCAGCTTCTTGTGACCTTGAAACCGACACAGACATTAATGTCCATGCTAACA	6880	7901	ACTGCAAGCTCTCTACCTTGGAGGAGGATGCCAGTTTAAACCCACCCACTCTGCANGA	7960
2181	AspProSerHisIleThrAlaGluThrAlaAlaArgLeuAlaArgGlySerProPro	2200	2541	SerLysTyrGlyPheGlyAlaLysGluValArgSerLeuSerGlyArgAlaValAsnHis	2560
6881	GATCCATCTCATATCAGCGGAGACCGCAGCGCGGCTTGGCGGGGGTCAACCCCA	6940	7961	TCCAAGTACGGGTTTGGGCTAAGGAGGTCGCGAGCTTGTCCGGGAGGCGGTTAACCC	8020
2201	SerGluAlaSerSerAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThr	2220	2561	IleLysSerValTrpLysAspLeuLeuGluAspSerGluThrProIleProThrIle	2580
6941	TCTGAGCGACGCTCTCAGCGAGCCAGTTATCGCGCCATCGCTGCGAGCCACTGCACC	7000	8021	ATCAAGTCCGTGTGGAGGAGCTCTCTGGAGGACTCACAACACCAATTCACCAACCATC	8080
2221	ThrHisGlyLysAlaTyrAspValAspMetValAspAlaAsnLeuPheMetGlyGlyAsp	2240	2581	MetAlaLysAsnGluValPheCysValAspProThrLysGlyGlyLysLysAlaAlaArg	2600
7001	ACCACGGCAGACCATGATGTGGACATGGTGGATGCCAACATTTATGGGGGGCAT	7060	8081	ATGGCCAAAATAGAGGTTCCTGCTGGAGCCCCCAAGAGGGGTAAAGAACGAGCTCGC	8140
2241	ValThrArgIleGluSerGlySerLysValValLeuAspSerLeuAspProMetVal	2260	2601	LeuIleValTyrProAspLeuGlyValArgValCysGluLysMetAlaLeuTyrAspIle	2620
7061	GTCACTCGATAGTCCGAGTCCAAAGTGTCTCTGGACTCTCTCGATCCATGGTTC	7120	8141	CITATCGTTTACCTGACCTCGCGCTCAGGCTCGCGAGAGATGGCCCTTTATGATGTC	8200
2261	GluGluArgSerAspLeuGluProSerIleProSerGluTyrMetLeuProLysLysArg	2280	2621	ThrGlnLysLeuProGlnAlaValMetGlyAlaSerTyrGlyPheGlnTyrSerProAla	2640
7121	GAAGAAGAGGGCGACCTCGAGCCTTCGATACCATCAGAGTACATGCTCCCCAAAAGAAG	7180	8201	ACACAAAATCTCCTCAGCGGTAAATGGGGGCTTCTTATGGATTCCAGTACTCCCGGCC	8260
2281	PheProAlaLeuProAlaTyrAlaArgProAspTyrAsnProProLeuValGluSer	2300	2641	GlnArgValGluPheLeuLeuLysAlaTyrAlaGluLysLysAspProMetGlyPheSer	2660
7181	TTTCCACACGCTTACCGCTTGGCGCGCCAGACTACACACCCGCTTGGGAATCA	7240	8261	CAGCGGTGGAGTTCTCTTGAAGGATCGCGGGAAGAAAGAACCTTATGGGTTCG	8320
2301	TrpLysArgProAspTyrGlnProAlaThrValAlaGlyCysAlaLeuProProArg	2320	2661	TyrAspThrArgCysPheAspSerThrValThrGluArgAspIleArgThrGluGluSer	2680
7241	CGGAAGAGCCAGCTATCAACCGCCACCTGTGTGGGGTGTCTCTCCCCCTTAAG	7300	8321	TATGATACCCGATGCTTTCACACACCGCTCACCGAGAGACATCAGGACTGAGGAGTCC	8380
2321	LysThrProThrProProArgArgArgArgThrValGlyLeuSerGluAspSerIle	2340	2681	IleTyrArgAlaCysSerLeuProGluGluAlaHisThrAlaIleHisSerLeuThrGlu	2700
7301	AGACCCCGACGCGCCCGCCAGAGCGCGGACAGTGGGTCTGAGTGAGACCATTA	7360	8381	ATATACCGGCGCTTCTCCCTCGCGAGGAGGCCACACATGCCATACACTCGCTAACTGAG	8440
2341	GlyAspAlaLeuGlnGlnLeuAlaIleLysSerPheGlyGlnProProProSerGlyAsp	2360	2701	ArgLeuTyrValGlyGlyProMetPheAsnSerLysGlyGlnThrCysGlyTyrArgArg	2720
7361	GGGGATGCCCTCAACAGCTGGCTATTAAGGCTTTGGCCAGCGCCCGCCCAAGCGGAGT	7420	8441	AGACTTTACGTGGAGGGCTTATTTAAACAGAGGGCCAGTCTCGCGGGTACAGCGCT	8500
2361	SerGlyLeuSerThrGlyAlaGlyAlaAlaAspSerGlySerGlnThrProAspGlu	2380	2721	CysArgAlaSerGlyValLeuThrThrSerMetGlyAsnThrIleThrCysTyrValLys	2740
7421	TCAGGCTGTCCACGGGGGGGCGCGCGGCTCGGAGTCCGACGCGCCCTGATGAG	7480	8501	TGCGCGCCAGTGGGGTGTCTACCATAGCATGGGGAACACCATCAGTGTCTAGCAGAAA	8560
2381	LeuAlaLeuSerGluThrGlySerIleSerSerMetProLeuGluGlyGluLeuGly	2400	2741	AlaLeuAlaAlaCysLysAlaAlaGlyIleIleAlaProThrMetLeuValCysGlyAsp	2760
7481	TCGGCTCTTCGGAGACAGGTTCCATCTCTCCATGCCCGCCCTCGAGGGGAGCCTGGA	7540	8561	GCCCTAGCGGCTGCAAGGCTGAGGATAGTTGGCGCCACCAATGCTGCTATGCGCGCAG	8620
2401	AspProAspLeuGluProGluGlnValGluProGlnProProGlnGlyValAla	2420	2761	AspLeuValValIleSerGluSerGlnGlyThrGluGluAspGluArgAsnLeuArgAla	2780
7541	GATCCAGACTGAGCGCTGAGCAGGTAGAGCGTCAACCTCCCCCGGGGGGGAAGTA	7600	8621	GACTTGGTGTCTATCTCAGAAAGCCAGGGGAGCCGAGGAGATGAGCGGACCTGAGAGCC	8680
2421	AlaProGlySerAspSerGlySerTrpSerThrCysSerGluGluAspSerValVal	2440	2781	PheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProArgProGluTyr	2800
7601	GCTCCGGCTCGGACTCGGGGCTCTGCTCACTTCTCGGAGGAGGAGCACTCCGCTG	7660	8681	TTCAAGGAGCTATGACCAAGGTATCTGCCCTCTCGCGGAGCCCGCCAGCGGATAT	8740
2441	CysCysSerMetSerTyrSerThrThrGlyAlaLeuIleThrProCysSerProGluGlu	2460	2801	AspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaLeuGlyProGlnGly	2820
7661	TGCTGCTCCATGCTATCTCTGGACCGGGGCTCTAATAACTCTTGTAGCCCCGGAAGAG	7720	8741	GACTTGGAGCTGATAACATCTTGTCTCAATGTCTGTGGCGTGGCGCCCAAGGC	8800
2461	GluLysLeuProIleAsnProLeuSerAsnSerLeuLeuArgTyrHisAsnLysValTyr	2480	2821	ArgArgArgTyrTyrLeuThrArgAspProThrThrProIleAlaArgAlaTyrGlu	2840
7721	GAAGAATTACCGATTAAACCTTTGACCAACTCGCTTTCGATACCAACCAAGGTGTAC	7780	8801	CGCCCGAGATATCTACTGACGAGACCTTACCTCCAAATCGCCGGGCTGCTCGGAA	8860
2481	CysThrThrLysSerAlaSerLeuArgAlaLysLysValThrPheAspArgMetGln	2500	2841	ThrValArgHisSerProValAsnSerTrpLeuGlyAsnIleIleGlnTyrAlaProThr	2860
7781	TGTACCAATCAAGAGCGCTCCTCAGGGCTTAAAGGTAACTTTTGTATGATGATGCAA	7840	8861	ACAGTTAGACACTCCCTCGCCANTTCTGGCTGGGAAACATCATTCAGTACGCCCCGACC	8920
2501	ValLeuAspSerTyrTyrAspSerValLeuLysAspIleLysLeuAlaLysVal	2520	2861	IleTyrAlaArgMetValLeuMetThrHisPhePheSerIleLeuMetAlaGlnAspThr	2880
			8921	ATATGGGCTCGCATGCTGATGACACACTTCTTCTCCATCTCTCATGGCCCAAGACACT	8980

991	SerProMetGluLysLysValIleValTrpGlyAlaGluThrAlaAlaCysGlyAspIle	1000
3281	AGUCGGAUGGAGAAAAAGAUCAUCUCUGCGGAGCGGAGACAGCUGCUGCGGGGAUUC	3340
1001	LeuHisGlyLeuProValSerAlaargLeuGlyArgGluValLeuLeuGlyProAlaAsp	1020
3341	UUACACAGCAUUCUCCUGUGCCGCCGCAUCUGCCGGAGGUCUCCUUGGCCACUGAU	3400
1021	GlyTyrThrSerLysGlyTrpSerLeuLeuAlaProIleThrAlaTyrAlaGlnGlnThr	1040
3401	GGCUAUAUCCCAAGGGGUGGAGUUCUCGCCCCCAUCACUGCUUAUGCCAGCAGACA	3460
1041	ArgGlyLeuLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGluGlnAla	1060
3461	CGCGCCUUUUUGGGCACCAUAGUGUGAGCAGCAUCAGCGGGCGCCACAAGACACAA	3520
1061	GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly	1080
3521	GGGGAGAUUCUAGGUCUUGCCACAGGUCACUAGUCCUUCUCCGGAACAACCAUCU	3580
1081	ValLeuTrpThrValTyrHisGlyValAlaGlyAsnLysThrLeuAlaGlySerArgGly	1100
3581	GUCUUAUGGACUGUCUACCAUGAGGAGCUGGCAACAAGACUCUANGCCGCGCAC	3640
1101	ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTyrProSerProGlyI	1120
3641	GUCAACAGAUUAUCCUCCAGUGCGAGCGGACUUAUGGGGUGGCCACAGCCGCCG	3700
1121	ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArgAsn	1140
3701	ACCAAAAUUUUGGAGCCUGGACGUGUGGAGCGGUCGACCUUAUACUUGGUCACG	3760
1141	AlaAspValIleProAlaArgArgGlyAspLysArgGlyValAlaLeuLeuSerProAs	1160
3761	CGUAUGUCAUCCCGGCUCCGAGAGACCGGGGAGCAAGCGAGGAGCGCUACUCCU	3820
1161	ProLeuSerThrLeuLysGlySerSerGlyGlyProValLeuCysProArgGlyHisAla	1180
3821	CCUUCUCCACUUGAAGGGUCCUCGGGGGGCCCGGUGCUCUGCCCCAGAGGCCAC	3880
1181	ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle	1200
3881	GUCGGGGUCUCCCGGGACGCGUGUGUCCCGGGCGGCGGCCAAGUCUAUAGAUAU	3940
1201	ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro	1220
3941	CCCGUUGAGACAUUGACUACGUCACUCGGUCCCCCACCUCUUGUGACAAACAGCAC	4000
1221	ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLys	1240
4001	CCUGCUGGCCCCAAACUUAUACAGGUCGGGUACUUAACUGCCCCGACUGGUAUG	4060
1241	SerThrLysValProValAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnPro	1260
4061	AGCACCAAGAUCCUUGUGCGUAGUCCGUCAGGGGUACAAAGUGUCUAGUGCUAAU	4120
1261	SerValAlaAlaThrLeuGlyPheGlyValaTyrLeuSerLysAlaHisGlyIleAsnPro	1280
4121	UCGGUGGUGGUGCCUCCUGGGGUUUGGGGCGUACUUGUCCAGGGCACAUUGGCAU	4180
1281	AsnIleArgThrGlyValArgThrValThrThrGlyAlaProIleThrTyrSerThrTyr	1300
4181	AACAUAUAGGACUGGGGUCAGGACUGUGACGACCGGGGGCCCAUACACGUACCAU	4240
1301	GlyLysPheLeuAlaAspGlyGlyCysAlaGlyAlaTyrAspIleIleIleCysAsp	1320
4241	GGCAAAAUUCCUGCGCGAUUGGGGGUGCGACGCGCGCCUUAUGACAUCAUCAU	4300
1321	GluCysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGlnAla	1340
4301	GAAGGCCAUGCCGGGACUUCUACCAUUCUGGCAUCGGGAACAGUCCUGGAUCAG	4360
1341	GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProProGlySerValThr	1360

4361	Db		GAGACAGCGGGGUCAGGCUAAUCUGAUCUGGCUACGCGUACGCGCCCGGGGUCAGUGACA	4420
1361	Qy		ThrProHisProAsnIleGluGluValAlaLeuClvGlnGluGlyGluIleProPheTyr	1380
4421	Db		ACCCCCCAACCAUAUGAGAGGUGGGCCUCUGGGCAGGAGGUGAGAUCCCUUCUAU	4480
1381	Qy		GlyArgAlaIleProLeuSerTyrIleLysGlyVArgHisLeuIlePheCysHisSer	1400
4481	Db		GGGAGGGCGAUUCCCUUGUAUAUACNAAGGGAGGAGACACUUGAUCUUCGCGACUCA	4540
1401	Qy		LysLysLysCysAspGluLeuAlaAlaLeuArgGlyMetGlyLeuAsnSerValAla	1420
4541	Db		NAAGAAAAAGUGAGCAGCUCUGCGCGCCUCUUGGGUUAUGGGCUUUAAGCAGUGGCA	4600
1421	Qy		TyrTyrArgGlyLeuAspValSerValIleProThrGlnGlyAspValValValAla	1440
4601	Db		UACUACAGAGGGGUGGAGCUCUCGUAUAUACCAACUAGGGAGACGUGGUGGUGCGCC	4660
1441	Qy		ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnVal	1460
4661	Db		ACCGACGCCCUCAUGACGGGGUUUACUGGAGACUUUGACUCCGUGAGCAGUCACGUA	4720
1461	Qy		AlaValThrGlnValValAspPheSerLeuAspProThrPheThrIleThrGlnIle	1480
4721	Db		CGGUCUCACAAAGUUGAGACUUCAGCUGGACCCCAUUCACCAUAAACACACAGACU	4780
1481	Qy		ValProGlnAspAlaValSerArgSerGlnArgArgGlyArgThrGlyArgLeu	1500
4781	Db		GUCCUCUACAGCGUGUCUACGUAUGCAGCGCCCGGGCCGACCGGGAGAGACUG	4840
1501	Qy		GlyIleTyrArgTyrValSerThrGlyGluArgAlaSerGlyMetPheAspSerValVal	1520
4841	Db		GGUUAUUAUGUAUGUUUCCACUGGUGAGCGACCCUACGGAUUGUUUGACGUGUAGUG	4900
1521	Qy		LeuCysGluCysTyrAspAlaGlyAlaAlaIleTyrGluLeuThrProSerGluThrThr	1540
4901	Db		CUCGCGAGUGUCACGAUGCAGGGGCGCGAUGUAUGAGCUCACACCGCGGAGACACC	4960
1541	Qy		ValArgLeuArgAlaTyrPheAsnThrProGlyLeuProValCysGlnAspHisLeuGlu	1560
4961	Db		GUCAGGUCUACAGACAUUUUACAACACACUGGUGUUGCUGUGGCCAAGACCAUCUUGAG	5020
1561	Qy		PheTrpGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThr	1580
5021	Db		UUUUGGGAGGAGUUUUCACCGGCCUACACAUAGUGCCCAUUCUUUCCCAACA	5080
1581	Qy		LysGlnSerGlyGluAsnPheAlaTyrLeuThrAlaTyrGlnAlaThrValCysAlaArg	1600
5081	Db		AAGCAUUCGGGGAAAAUUUCGUAUCUUAACAGCCUACAGGCUACAGUGUGCGCUAGG	5140
1601	Qy		AlaLysAlaProProSerTrpAspValMetTrpLysCysLeuThrArgLeuLysPro	1620
5141	Db		GCCAAAGCCCCCCCCGUGGAGCUGUAGUGGAGUGUUGACUCGACUACAAGCCC	5200
1621	Qy		ThrLeuValGlyProThrProLeuLeuTyrArgLeuGlySerValThrAsnGluValThr	1640
5201	Db		ACACUGUGGGCCCCACACCUCCUGUACCGCUUGGCGUCUUGUUAUCCACAGAGUACCC	5260
1641	Qy		LeuThrHisProValThrLysTyrIleAlaThrCysMetGlnAlaAspLeuGluValMet	1660
5261	Db		CUCACGCAUCCUGAGCGAAUAUACUCCGCCACCUAGUGCAAGCCGACCUUAGGUCAUG	5320
1661	Qy		ThrSerThrTrpValLeuAlaGlyGlyValLeuAlaValAlaIleAlaIleTyrCysLeuAla	1680
5321	Db		ACCAGCACGUGGUGUCUUAAGCUGGGGGGUCUUGGGCGCGCGCGGUGACUCUGCGCG	5380
1681	Qy		ThrGlyCysValCysIleIleGlyArgLeuHisIleAsnGlnArgAlaValAlaPro	1700
5381	Db		ACCGGGUGUGUUUGCAUCAUCGCGCCUUGGACGUAUACCGAGCGCGGUGGUGACCG	5440
1701	Qy		AspLysGluValLeuTyrGluAlaPheAspGluMetGluGluCysAlaSerArgAlaAla	1720

Qy	2081	AsnPhelysValAlaIleTTPArgValAlaAlaSerGluTyrAlaGluValThrGlnHis	2100	Qy	2441	CysCysSerMetSerTyrSerTrpThrGlyAlaLeuIleThrProCysSerProGluGlu	2460
Db	6581	AACTUUAAGAUCCGCAUCUGAGGUGGCGCCUCACAGAUACGCGAGUGAGCGACAC	6640	Db	7661	UGCUGCUCCAUUGCAUACUCCUGGACCGGGGCGUCUAAUAACUCCUUGUAGCCCGAAGAG	7720
Qy	2101	GlySerTyrHisTyrIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeu	2120	Qy	2461	GluLysLeuProIleAsnProLeuSerAsnSerLeuLeuArgTyrHisAsnLysValTyr	2480
Db	6641	GGGUCUACACACUACAUACAGGACUACACUAGAUAAUUCUUGAAUUCUUGCCAAUA	6700	Db	7721	GAAGAUGUCCAAUUGGCGCCUUGAGCAACUCCUUGGUGCGAUUACCAACAGGUGUAC	7780
Qy	2121	ProSerProGluPhePheSerTTPValAspGlyValGlnIleHisArgPheAlaProThr	2140	Qy	2481	CysThrThrThrLysSerAlaSerLeuArgAlaLysLysValThrPheAspArgMetGln	2500
Db	6701	CCUUCUCCAGAUUUUUUCCUGGGUGGACGAGUGCGAUUCCAUAGGUUUUGCCCAUA	6760	Db	7781	UGUACCAACAAGAGCGCCCUUUAAGGGCUAANAAGUAACUUUUGUAAGGAUGCAA	7840
Qy	2141	ProLysProPhePheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal	2160	Qy	2501	ValLeuAspSerTyrTyrAspSerValLeuLysAspLysLeuAlaLalaSerLysVal	2520
Db	6761	CCGAAGCGUUUUUUCGGGAUGAGGUUCUUGUUGGUGGUUUAUUUUGUGUC	6820	Db	7841	GGCGUCGACGCUCAUUAUGACUCCAGCUUUGAAGGACAUUAAGCAGCGGCCUCCAAAGGUC	7900
Qy	2161	GlySerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThr	2180	Qy	2521	ThrAlaArgLeuLeuMetGluAlaCysGlnLeuThrProProHisSerAlaArg	2540
Db	6821	GGGUCUCCAGUCCUUGCGAUCCUGAACUGACACAGACGUUUGACUGUCCUUAACA	6880	Db	7901	ACCGCAAGGCUUCUUAUAGAGGAGGCGCCAGUUAACUCCACCCACUCUGCAAGA	7960
Qy	2181	AspProSerHisIleThrAlaGluThrAlaAlaArgArgLeuAlaArgGlySerProPro	2200	Qy	2541	SerLysTyrGlyPheGlyAlaLysGluValArgSerLeuSerGlyArgAlaValAsnHis	2560
Db	6881	GACCAUCCCAUUAUACGCGGAGACUGCAGCGCGGUUGGACCGGGGUGCACCCCG	6940	Db	7961	UCCAGUAUGGUGUUGGGCUAAGAGGUGCCGAGCUUGUCCGGGAGAGCCGUUAAACC	8020
Qy	2201	SerGluAlaSerSerAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThr	2220	Qy	2561	IleLysSerValTrpLysAspLeuLeuGluAspSerGluThrProIleProThrThrIle	2580
Db	6941	UCCGAGGACAGUCCUCCAGCGAGCCAGCUAUCGCGACCAUCUGCGAGCCACUCCGACC	7000	Db	8021	AUCAAGUCCUGUGGAGGACCUCCUGGAAGACACACACAAACCAUUAUCCACCAUC	8080
Qy	2221	ThrHisGlyLysAlaTyrAspValAspMetValAspAlaAsnLeuPheMetGlyLysAsp	2240	Qy	2581	MetAlaLysAsnGluValPheCysValAspProThrLysGlyGlyLysLysAlaAlaArg	2600
Db	7001	ACCCAGCGCAAGCGCUAUGAUGGACUAGUGGACUAGUGGACCAUCCUUGUUGGCGCAU	7060	Db	8081	AUGGCCAAAUUAGAGGUGUUCUGCGUGGACCCACCAAGGGGGUAGAAAGCAGCUCGC	8140
Qy	2241	ValThrArgIleGluSerGlySerLysValValValLeuAspSerLeuAspProMetVal	2260	Qy	2601	LeuIleValTyrProAspLeuGlyValArgValCysGluLysMetAlaLeuTyrAspIle	2620
Db	7061	GUGACCGGAGUAGUGUCUGAGUCCAAAGUGGUGUUGUGGACUCCUGGACCCCAUUGGUC	7120	Db	8141	CUUAUCGUUUUACCCUGACCCUGCGGUGGUGGUGGAGAAUUGGCCCUUUAUGAUAUC	8200
Qy	2261	GluGluArgSerAspLeuGluProSerIleProSerGluTyrMetLeuProLysLysArg	2280	Qy	2621	ThrGlnLysLeuProGlnAlaValMetGlyAlaSerTyrGlyPheGlnTyrSerProAla	2640
Db	7121	GAAGAAGGAGCGAGCCUUGAGCCUUCGAUACCAUCCGAUUAUUGUCCCAAGAGAGA	7180	Db	8201	ACACAAAGCUUCCUCCAGCGGUGAUGGGGCGUUCUUAUGGAUCCAGUACUCCCGCU	8260
Qy	2281	PheProProAlaLeuProAlaTTPAlaArgProAspTyrAsnProProLeuValGluSer	2300	Qy	2641	GlnArgValGluPheLeuLeuLysAlaTTPAlaGluLysLysAspProMetGlyPheSer	2660
Db	7181	UUCGCCACCGCUUACCGGUGGCGACGCGCCUGAUUACACCCAGCCGCGUUGUGAAUCG	7240	Db	8261	CAGCGGUGGAGUUCUUCUUGAAGCAGUGGGCGGAAAGAAAGACCCUUGGGUUUUCG	8320
Qy	2301	TrpLysArgProAspTyrGlnProAlaThrValAlaGlyCysAlaLeuProProArg	2320	Qy	2661	TyrAspThrArgCysPheAspSerThrValThrGluArgAspIleArgThrGluGluSer	2680
Db	7241	UGGAAGAGGCGCAUUAACCAACCGCCACUUGUUGCGGCGUGGCUCCUCCUCCUUAAG	7300	Db	8321	UAUGAUCCCGAUGUCCUUGACUACCCGUCACUGAGAGAGACAUAGGACUGAGGAGUCC	8380
Qy	2321	LysThrProThrProProProArgArgArgThrValGlyLeuSerGluAspSerIle	2340	Qy	2681	IleTyrArgAlaCysSerLeuProGluGluAlaHisThrAlaIleHisSerLeuThrGlu	2700
Db	7301	AAAACCCGACGCGUCCUCCAGGAGAGCGCGGACAGUGGUGUGAGAGAGUCCUUA	7360	Db	8381	AUAUAUCCGGGUUUGUCCUUGCCCGAGGAGGCGCCACACUCCCAUACACUACUGAG	8440
Qy	2341	GlyAspAlaLeuGlnLeuAlaIleLysSerPheGlyGlnProProProSerGlyAsp	2360	Qy	2701	ArgLeuTyrValGlyGlyProMetPheAsnSerLysGlyGlnThrCysGlyTyrArgArg	2720
Db	7361	GCAGAUGCCUUAACAACAGUGGCGCAUUAAGUCCUUGGCCAGCCUCCUCCCAAGCGCAU	7420	Db	8441	AGACUUUACUGGAGGAGGCGCCAUUUAACAGCAAGGCGCAGACCUCCGCGUACAGCGCU	8500
Qy	2361	SerGlyLeuSerThrGlyAlaGlyAlaAlaAspSerGlySerGlnThrProProAspGlu	2380	Qy	2721	CysArgAlaSerGlyValLeuThrThrSerMetGlyAsnThrIleThrCysTyrValLys	2740
Db	7421	UCAGGCGUUUCCACGGGGCGGACGCGAGCGAUUCCGGACUGGAGCGCGCCUCCGAGAG	7480	Db	8501	UGCCGCGCCAGCGGGGUGUUAUCCACUAGAUUGGGAAACCAUACAUAGCUUUGUAGAA	8560
Qy	2381	LeuAlaLeuSerGluThrGlySerIleSerSerMetProProLeuGluGlyLeuGly	2400	Qy	2741	AlaLeuAlaAlaCysLysAlaAlaGlyIleIleAlaProThrMetLeuValCysGlyAsp	2760
Db	7481	UUGGCCUUUUGGAGACAGGUUUAUCCUCCUAGGCCUCCUCCUCCGAGGGGAGCCUGGA	7540	Db	8561	GCCUAGCGCCUUAAGGUGUGGAGGUAUUAUUGGCCACAAUUGGUGUAGUCCGCGAU	8620
Qy	2401	AspProAspLeuGluProGluGlnValGluProGlnProProProGlnGlyValAla	2420	Qy	2761	AspLeuValValIleSerGluSerGlnGlyThrGluGluAspGluArgAsnLeuArgAla	2780
Db	7541	GAUCCAGCUUGGAGCCUGAGCAGGAGUAGAGUACUACUCCUCCCGAGGGGGGUGGUA	7600	Db	8621	GACUUGUUGUACUUCACAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGC	8680
Qy	2421	AlaProGlySerAspSerGlySerTrpSerThrCysSerGluGluAspSerValVal	2440	Qy	2781	PheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProArgProGluTyr	2800
Db	7601	ACCCCGGCGUACGCGGUGUUGGUGUACUUGUCCUGAGGAGGACGACUCCGUGUG	7660	Db	8681	UUCACGGAGGCUAUGACCAAGGUUUCUCCUCCUUGGAGUCCUCCCGAGCGGAAU	8740
				Qy	2801	AspLeuGluLeuIleThrSerCysSerAsnValSerValAlaLeuGlyProGlnGly	2820

```

8741 GACCGAGCUGAUAACUUCUGCCUCAAUUGUGUGUGGCGCCACCAAGGC 8800
2821 ArgArgGlyTyrLeuThrArgAspProThrProIleAlaArgAlaThrGlu 2840
8801 CGCCGAGAUACUACUGACAGACACCCUACCAUCCGCGGCGUGCCUGGAA 8860
2841 ThrValArgHisSerProValAsnSerTrpLeuGlyAsnIleIleGlnTyrAlaProThr 2860
8861 ACAGUAGACACUCCUUGCAUUAUGGCUAGGAAACUACUACUAGUACGCCCCARCC 8920
2861 IleTrpAlaArgMetValLeuMetThrHisPheSerIleLeuMetAlaGlnAspThr 2880
8921 AUAUGGCGCUGAUGGUGUGAUGACACAUUCUUCUACUUGGCGCCCAAGAUACU 8980
2881 LeuAspGlnAsnLeuLeuPheGluMetTyrGlyAlaValTyrSerValSerProLeuAsp 2900
8981 CUGACCAAGACUACUACUUGAUGUACGAGCGGUGUACUCCUGAGUCCUUGGAC 9040
2901 LeuProAlaIleIleGluArgLeuHisGlyLeuAspAlaPheSerLeuHisThrTyrThr 2920
9041 CUCCAGCAUAUUAAGAAUGUACACGCGCUUGACGCUUCUUCUGACACAUACU 9100
2921 ProHisGluLeuThrArgValAlaSerAlaLeuArgGlyLeuGlyAlaProProLeuArg 2940
9101 CCCCAAGACUGACACGCGGUGUACUACGCGUACGCGGCGCCACCCUCAGA 9160
2941 AlaTrpIleSerArgAlaArgAlaValAlaSerLeuIleSerArgGlyGlyArgAla 2960
9161 GCGUGAAGACGCGGCGACGUGACGUGACGCGGCGUCCUACUCCGCGGCGGAGCG 9220
2961 AlaValCysGlyArgTyrLeuPheAsnTrpAlaValIleThrIleLeuLeuThrPro 2980
9221 GCGUUGCGCGCAUAUUCUUCUACUGGCGGUGAAGAACCAAGCUCAACUACUCCA 9280
2981 LeuProGluAlaArgLeuLeuAspLeuSerSerTrpPheThrValGlyAlaGlyGly 3000
9281 UUGCGGAAGCGCGGCGUCCUGAUUAUACUGGUGUACUUGCGGCGCGGCGGCG 9340
3001 AspileTyrHisSerValSerArgAlaArgProArgLeuLeuLeuPheGlyLeuLeu 3020
9341 GACAUUAUACACGUGUGCGUGCGCGACCGCGCUUUAUUCUCCUUGGCGUACUCCA 9400
3021 LeuPheValGlyValGlyLeuPheLeuLeuProAlaArg 3033
9401 CUUUUUGAGGGUGGUGGCUUUUUCUACUCCCGCGCGG 9439

```

```

RESULT 2
US-07-925-695-2
Sequence 2, Application US/07925695
Patent No. 5428145
GENERAL INFORMATION:
APPLICANT: OKAMOTO, Hiroaki
APPLICANT: NAKAMURA, Tetsuo
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOME,
TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, ANTIGEN, ANTIBODY AND
TITLE OF INVENTION: DETECTION METHODS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Beveridge, DeGrandi, Weilacher & Young
STREET: 1850 M Street, N.W., Suite 800
CITY: Washington
STATE: D.C.
COUNTRY: US
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07925,695
FILING DATE: 19920807

```

```

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 287402/91
FILING DATE: 09-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 360441/91
FILING DATE: 05-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Weilacher, Robert G.
REGISTRATION NUMBER: 20,531
REFERENCE/DOCKET NUMBER: 06/87-48009
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2811
TELEFAX: (202) 659-1462
TELEX: WUI 64470
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 9589 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
US-07-925-695-2
Alignment Scores:
Pred. No.: 0 Length: 9589
Score: 15824.00 Matches: 2966
Percent Similarity: 98.78% Conservative: 30
Best Local Similarity: 97.79% Mismatches: 37
Query Match: 98.16% Indels: 0
DB: 1 Gaps: 0

```

```

US-09-980-559-2 (1-3033) x US-07-925-695-2 (1-9589)
QY 1 MetSerThrAsnProIleProGlnArgIleThrIleAsnThrAsnArgProGln 20
DB 341 ATGAGCACAAATCTTAAACCTCAAGAGAAACCAAAAGAAACCAACCGTCCGCCAACAA 400
QY 21 AspValIlePheProGlyGlyGlnIleValGlyGlyValIleValIleLeuProArgArg 40
DB 401 GAGCTTAACTTCCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 460
QY 41 GlyProArgLeuGlyValArgAlaThrArgIleThrSerGluArgSerGlnProArgGly 60
DB 461 GCGCCCGAGCTTGGGTGTGGCGCGCGACAGAGAGACTTCGAGCGGTCGCCAGCACCGTGA 520
QY 61 ArgArgGlnProIleProIleProIleProIleProIleProIleProIleProIleProIle 80
DB 521 AGCGCCAGCCCATCCCTAAGGATCGGCGCTCCACTGGCAAAATCCTGGGGAHAACAGGA 580
QY 81 TyrProTrpProLeuTyrGlyAsnGluGlyLeuGlyTyrAlaGlyTrpLeuLeuSerPro 100
DB 581 TACCCCTGGCCCTATACGGGATGAGGAGCTCGGCTGGGAGGATGGCTCCTGTCTCCCC 640
QY 101 ArgGlySerArgProSerTrpGlyProAsnAspProArgHisArgSerArgAsnValGly 120
DB 641 CGAGGTTCCTCGTCCCTCTTGGGCGCCCAATGACCCCGGCGCATAGGTCCGCAACGTGGGT 700
QY 121 LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetClyTyrIleProValVal 140
DB 701 AAGGTCAATGATACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT 760
QY 141 GlyAlaProLeuGlyGlyValAlaArgAlaLeuAlaHisGlyValArgValLeuGluAsp 160
DB 761 GCGCCCGCGCTCGGCGGCGCTCGGCGAGCTCTCGGCGAGTGGCGTGGAGTCTCTGGAGGAC 820
QY 161 GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla 180
DB 821 GGGGTTAAATTTGCAACAGAGAACTTACCCGGTGTCTCTCTTTCTATCTCTTTCTGTGGCC 880
QY 181 LeuLeuSerCysIleThrThrProValSerAlaAlaGluValIleSerThrGly 200
DB 881 CTGCTGTCTGTGCATCACCACCGCGGTCTCCGCTGCGGAAAGTGAAGAACATCAGTACCGGC 940

```

Qy	201	TyrMetValThrAsnAspCysThrAsnAspSerIleThrTrpGlnLeuGlnAlaVal	220
Db	941	TACATGGTGACCAACGACTGACCAATGATAGCATTACCTGGCACTCCAGGCTGTGTC	1000
Qy	221	LeuHisValProGlyCysValProCysGluValGlyValAsnAlaSerGlnCysTrpIle	240
Db	1001	CTCACGTCCTCCGGTGGTCCGTCGGGAAAGTGGGGAATACATCTCGTCTGGATA	1060
Qy	241	ProValSerProAsnValAlaValGlnArgProGlyAlaLeuThrGlnGlyLeuArgThr	260
Db	1061	CCGGTCTCCAGGAATGGCGGTGGCAGCAGCCGGCGCCCTCACGCGAGGCTTACGAGC	1120
Qy	261	HisIleAspMetValValMetSerAlaThrLeuCysSerAlaLeuTyrValGlyAspLeu	280
Db	1121	CACATTGACATGGTGTGTGATGTCGCGCACCTCTGCTCCGCTCTTTACGTGGGGACCTC	1180
Qy	281	CysGlyGlyValMetLeuAlaAlaGlnMetPheIleValSerProGlnHisHisTrpPhe	300
Db	1181	TGCGGTGGGTGATGCTTGGAGCCAGATGTTTCAATGTCTCGCCACAGCACCACTGGTTT	1240
Qy	301	ValGlnAspCysAsnCysSerIleTyrProGlyThrIleThrGlyHisArgMetAlaTrp	320
Db	1241	GTGCAAGACTGCAATGCTCTCCATCTACCTGTGTACCATCACTGACACCCGATGGCGTG	1300
Qy	321	AspMetMetMetAsnTrpSerProThrAlaThrMetIleLeuAlaTyrAlaMetArgVal	340
Db	1301	GACATGATGATGAATGCTGCGCCACCGCTPACCATGATCTGCGTACCGGATCGCGCTC	1360
Qy	341	ProGluValIleIleAspIleIleSerGlyAlaHisTrpGlyValMetPheGlyLeuAla	360
Db	1361	CCGAGGTCACTATGACATCATTTGGCGGGCTCATTTGGGGCGTCAATTTGGGCTTAGCC	1420
Qy	361	TyrPheSerMetGlnGlyAlaTrpAlaLysValValIleLeuLeuAlaAlaGly	380
Db	1421	TACTTCTATGACGGAGCGTGGCAAAAGTCTGTGTCACTTTTGTGCGCGCGG	1480
Qy	381	ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr	400
Db	1481	GTGGACGCGCAACCCCACTACCGTTGGGGTCTTACCGCGCATAAACGCCAGGACCTCACC	1540
Qy	401	SerLeuPheAspMetGlyProArgGlnIleValIleGlnLeuValAsnThrAsnGlySerTrp	420
Db	1541	GGCATGTTCTCCCTTGTGTCAGCGCAAAATCCAGCTCATCAACCAACCAATGGCAGTTGG	1600
Qy	421	HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer	440
Db	1601	CACATCAACCGCACCGCCCTGAATGCAATGACTCTTTGACACCGGCTTCTCGGTCA	1660
Qy	441	LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg	460
Db	1661	CTGTTCTACACCCACAGCTTCAACTGTCAGGATGTCCGAAACGCAATGTCGCGCTGCCG	1720
Qy	461	SerIleGluAlaPheArgValGlyTrpGlyAlaLeuGlnTyrGluAspAsnValThrAsn	480
Db	1721	AGTATCGAGGCTTTTGGGTGGATGGGCGCCTTACAATATAGACAAATGTCACCAAT	1780
Qy	481	ProGluAspMetArgProTyrCysTrpHisTyrProProArgGlnCysGlyValValSer	500
Db	1781	CCAGAGATATGAGCGTATTCGTGGCACTACCCACCAAGACAGTGTGTGTAGTCTCC	1840
Qy	501	AlaLysThrValCysGlyProValTyrCysPheThrProSerProValValGlyThr	520
Db	1841	GCGAGCTCTGTGTGGCGCCAGTACTGTTTCAACCCCGACGACGCTAGTAGTGGGTACG	1900
Qy	521	ThrAspArgLeuGlyAlaProThrTyrThrTrpGlyGluAsnGlnThrAspValPheLeu	540
Db	1901	ACCGATAGACTTGGAGCGCCACTTACACGTGGGGGAGAAATGAGACAGATGCTTCCCTA	1960
Qy	541	LeuAsnSerThrArgProProLeuGlySerTrpPheGlyCysThrTrpMetAsnSerSer	560
Db	1961	TTGACACGACTCGACCAACCGAGGGGTATGGTTGGCTGACAGTGGATGAATCACT	2020
Qy	561	GlyTyrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer	580
Qy	2021	GGCTACCAACAGACTTGGCGGCACACCCCTCCGCAATTAGAGCTGACTTCAATGCCAGC	2080
Qy	581	ThrAspLeuLeuCysProThrAspCysPheArgLysHisProAspThrThrTyrLeuLys	600
Db	2081	ATGGACTTGTGTGCCCCACGAGACTGTTTTAGAGACATCTCTGATACCACTACATCAAA	2140
Qy	601	CysGlySerGlyProTrpLeuThrProArgCysLeuIleAspTyrProTyrArgLeuTrp	620
Db	2141	TGTGGCTCTGGGCCCTGGCTCACGCCAAGTGCCTGATCGACTTACCCCTACAGGCTCTGG	2200
Qy	621	HisTyrProCysThrValAsnTyrThrIlePheLysIleArgMetTyrValGlyGlyVal	640
Db	2201	CATTACCCCTGCACAGTTAACTATACCATCTTCAAAATAGGATGTATGTGGGGGGGTC	2260
Qy	641	GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp	660
Db	2261	GAGCACAGGCTCAGGCTCGTGCAATTTCACTCGTGGGGATCGTTGCAACTTGGAGGAC	2320
Qy	661	ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGluTrpAlaIleLeuPro	680
Db	2321	AGAGACAGAACTCACTGTCTCTCTTGTGCTCACTCCACACGAGTGGGCCATTTTACCT	2380
Qy	681	CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisLeuHisGlnAsnIle	700
Db	2381	TGCACCTACTCGGACCTGCGCCCTTGTGCACTGGTCTTCTCCACCTCCACCAAAACATC	2440
Qy	701	ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrLysTyrIleValArgTrp	720
Db	2441	GTGGAGCTCAATTCATGTATGGCTATCACTGCTCTCAAAATACATCGTCCGATGG	2500
Qy	721	GluTrpValIleLeuLeuPheLeuLeuAlaAspAlaArgValCysAlaCysLeuTrp	740
Db	2501	GAGTGGGTAGTACTTATTTCTGCTTGTAGCGAGCGCGAGGTTTGGCTTATGG	2560
Qy	741	MetLeuIleLeuLeuGlyGlnAlaGluAlaAlaLeuGluLysLeuValIleLeuHisAla	760
Db	2561	ATGCTCATCTTGTGGCGCAGCCGAGCAGCACTAGAGAAGTGTGCTCTTGGACGCT	2620
Qy	761	AlaSerAlaAlaSerCysAsnGlyPheLeuTyrPheValIlePhePheValAlaAlaTrp	780
Db	2621	GGAGCGCAGCTAGCTGCAATGGCTTCTTATATTGTCTTCTTCTTCTGCTGCTGCTGG	2680
Qy	781	TyrIleLysGlyArgValValProLeuAlaThrTyrSerLeuThrGlyLeuTrpSerPhe	800
Db	2681	TACATCAAGGTCGGGTAGTCCCTTGTGCTACTTATTCCTCACTGCGCTATGCTCTTT	2740
Qy	801	SerLeuLeuLeuAlaLeuProGlnGlnAlaTyrAlaTyrAspAlaSerValHisGly	820
Db	2741	GGCTACTGCTCTAGCATTTGCCCAACAGGCTTATGCTTATGACGCATCTGTACATGGT	2800
Qy	821	GlnIleGlyAlaAlaLeuLeuValMetIleThrLeuPheThrLeuThrProGlyTyrLys	840
Db	2801	CAGATAGAGCAGCTCTGTGTGTGCTGCTATCTCTTATCACTCACCCCGGATATAAG	2860
Qy	841	ThrLeuLeuSerArgPheLeuTrpTrpLeuCysTyrLeuLeuThrLeuGlyGluAlaMet	860
Db	2861	ACCTTCTCAGCGGTTCTGTGGTGTGTGCTATCTTCTGACCCCTGGCGGAAGCTATG	2920
Qy	861	ValGlnGluTrpAlaProProMetGlnValArgGlyGlyArgAspGlyIleIleTrpAla	880
Db	2921	GTCCAGGAGTGGGCACCATATGAGGTCGCGGTGGCGGTGATGGGATCATATGGGCG	2980
Qy	881	ValAlaIlePheTyrProGlyValValPheAspIleThrLysTrpLeuLeuAlaValLeu	900
Db	2981	GTGCGCATATCTGCGCGGGTGTGGTGTGTGATCAACCAAGTGGCTCTTGGCGGTGCTT	3040
Qy	901	GlyProAlaTyrLeuLeuLysGlyAlaLeuThrArgValProTyrPheValArgAlaHis	920
Db	3041	GGGCTGCTTATCTCTAAAGGTGCTTTGACGCGGTGCGCTACTTCTCGTCAAGGCTCAC	3100
Qy	921	AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyArgTyrValGlnMet	940

3101 GCTCTACTAGGATGTGCACCATGTAAGCATCTCGCGGGGTAGGTACGTCCAGATG 3160
 941 AlaLeuAlaLeuGlyArgTTPThrGlyThrTyrLeuThrPheMet 960
 3161 GTGCTACTAGCCCTTGGCAGGTGGACATGACATCTATACCACTTCACCCCTATG 3220
 961 SerAspTTPAlaAlaSerGlyLeuArgAspLeuAlaValGluProIlePhe 980
 3221 TCGGATGGGCTGCTAATGCGCTGGGACTTGGCGGTGGCGGTATCATCTTC 3280
 981 SerProMetGluLysLysValIleValTTPGlyAlaGluThrAlaAlaCysGlyAspIle 1000
 3281 AGTCGATGGAGAAAAGTCACTGCTGGGAGCGGAGACGTGCTTGGCGGGATATC 3340
 1001 LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp 1020
 3341 TTACACGGACTTCCCGTGTCCCGGACTTGGCGGGAGGTCTCTCTGGCCCGAGCTGAT 3400
 1021 GlyTyrThrSerLysGlyTTPSerLeuLeuAlaProIleThrAlaTyrAlaGlnGlnThr 1040
 3401 GGCTATACCTTCCAAAGGGTGGAGTCTTCTCGCCCCCATCACTGTTATGCCGAGCAGACA 3460
 1041 ArgGlyLeuLeuGlyThrIleValSerMetThrGlyArgAspLysThrGluGlnAla 1060
 3461 CGCGGCTTTTGGGACCATAGTGTGAGCATGACGGGGCGGACAAAGACAGAACAGGCC 3520
 1061 GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly 1080
 3521 GGGGAGATTCAGGTCCTCCAGCGTCACAGTCTTCTCGGAACAACCATCTCGGG 3580
 1081 ValLeuTTPThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro 1100
 3581 GTCTTATGAGTGTCTACCATGAGCTGGCACAAGACTCTAGCCGGCTCACGGGTCCG 3640
 1101 ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTTPProSerProProGly 1120
 3641 GTCACACAGATGTACTCCAGTGTGAGGGGACTTAGTGGGGTGGCCAGCCCGCCGGG 3700
 1121 ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArgAsn 1140
 3701 ACCAAATCTTTGGAGCCGTGCAGTGTGGAGCGGTGACCTATACCTTGGTTCACGGCAAC 3760
 1141 AlaAspValIleProAlaArgArgGlyAspLysArgGlyAlaLeuLeuSerProArg 1160
 3761 GCTGATGTCTATCCCGTTCGAAGACGGGGGACAGCGAGAGCGTACTCTCCCGAGA 3820
 1161 ProLeuSerThrLeuLysGlySerSerGlyClyProValLeuCysProArgGlyHisAla 1180
 3821 CCTCTTCCACCTTGAAGGGGTCTCTCGGGGGCCCGGTGCTCTGCCCGACAGAGCCACGCT 3880
 1181 ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle 1200
 3881 GTCCGGGTCTTCCGGGCGAGCGGTGTCTCCGGGGCGTGGCCAGTCCATAGATTATTC 3940
 1201 ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro 1220
 3941 CCCGTTGAGACACTTGACATCGTCACTCGGTCCCGCCACCTTGTAGTACACAGCACACCA 4000
 1221 ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLys 1240
 4001 CCTGCTGTGCCCAAACTTATCAGTGGGTACTTACATGCCCGGACTGCTGGTAGTGGAAAG 4060
 1241 SerThrLysValProValAlaTyrAlaAlaGluClyTyrLysValLeuValLeuAsnPro 1260
 4061 AGCAACAAAGTCCCTGTCCGGGTATGCGGCTCAGGGGTACAAAGTGTAGTGTAAATCCC 4120
 1261 SerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerLysAlaHisGlyIleAsnPro 1280
 4121 TCGGTGGCTGCCACCTCGGGGTGGGGGTGAGTGTCTGTCGAAGGACATGGCATCAATCCC 4180
 1281 AsnIleArgThrGlyValArgThrValThrThrGlyAlaProIleThrTyrSerThrTyr 1300
 4181 AACATTAGGACTGGGGTCAGGACTGTGACGACCGGGCGGCCCATCAGGTACTCCACATAT 4240

1301 GlyLysPheLeuAlaAspGlyGlyCysAlaGlyGlyAlaTyrAspIleIleCysAsp 1320
 4241 GGCAAATCTCTCGCGATGGGGCTTGGCAGCGGGCTATGACATCATCATATCGCAT 4300
 1321 GluCysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGlnAla 1340
 4301 GAATGCCATCGCGTGGACTTACCACTTCTCGGCATCGGAACAGTCTCTCGATCAAGCA 4360
 1341 GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProProGlySerValThr 1360
 4361 GAGACAGCGCGGTGAGCTAATGTACTGGCTACCGCTTACCGCCCGGGGTGAGTGACA 4420
 1361 ThrProHisProAsnIleGluValAlaLeuGlyGlnGluGlyGluIleProPheTyr 1380
 4421 ACCCCCAACCAACATAGAGAGGTGGCCCTCGGGCAGGAGGTGAGATCCCTCTCTAT 4480
 1381 GlyArgAlaIleProLeuSerTyrIleLysGlyArgHisLeuIlePheCysHisSer 1400
 4481 GGGAGGGCGATTCCCTGTCTATACATCAAGGGAGGAAGACATCTGATCTTCTGCCACTCA 4540
 1401 LysLysLysCysAspGluLeuAlaAlaLeuArgGlyMetGlyLeuAsnSerValAla 1420
 4541 AAGAAAAGTGTGACGAGCTCGCGGGCTTTCGGGGTATGGGCTTGAACCGAGTGGCA 4600
 1421 TyrTyrArgGlyLeuAspValSerValIleProThrGlnGlyAspValValValAla 1440
 4601 TACTACAGAGGCTGGAGCTCTCCGTATATCACTCAGGGAGAGCTAGTGGTCTCGCC 4660
 1441 ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnVal 1460
 4661 ACCGAGCCCTCATGCGGGTCTTACTGGAGACTTTCGCTGATGATGATGATGATGATG 4720
 1461 AlaValThrGlnValValAspPheSerLeuAspProThrPheThrIleThrGlnIle 1480
 4721 GCGGTCACTCAAGTGTAGACTTTCAGTGGAGCCCACTTCCATTAACCAACACAGACT 4780
 1481 ValProGlnAspAlaValSerArgSerGlnArgArgGlyArgThrGlyArgGlyArgLeu 1500
 4781 GTCCCTCAAGACGCTGTCTCAGTAGCAGCGCGGGGCGCACGGGCGAGGAGAGCTG 4840
 1501 GlyIleTyrArgTyrValSerThrGlyGluArgAlaSerGlyMetPheAspSerValVal 1520
 4841 GGTATTTATAGGTATGTTTCCACTGGTGGAGCGGCTCAGGAATGTTTTCACAGTGTAGTG 4900
 1521 LeuCysGluCysTyrAspAlaGlyAlaAlaTyrTyrGluLeuThrProSerGluThrThr 1540
 4901 CTCTCGAGTGTCTACATGCGGGCGGCTATGAGTGTATGAGTGTATGAGTGTATGAGTGT 4960
 1541 ValArgLeuArgAlaTyrPheAsnThrProGlyLeuProValCysGlnAspHisLeuGlu 1560
 4961 GTCAGGCTCAGAGCATATTTCAACACACTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGT 5020
 1561 PheTTPGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThr 1580
 5021 TTTTGGGAGCGAGTTTTCACCGGCTCACACATAGATGCCACTTCTCTTCCCAACA 5080
 1581 LysGlnSerGlyGluAsnPheAlaTyrLeuThrAlaTyrGlnAlaThrValCysAlaArg 1600
 5081 AAGCAATCGGGGAAATTTCCGCATATTAAACAGCTTACAGGCTACAGTGTGGGTGAGT 5140
 1601 AlaLysAlaProProProSerTTPAspValMetTTPLysCysLeuThrArgLeuLysPro 1620
 5141 GGCNAAGCCCCCCCCCTCTGGGAGCTCATGTGGAAGTGTGTGACTCGACTCAAGCCC 5200
 1621 ThrLeuValGlyProThrProLeuLeuTyrArgLeuGlySerValThrAsnGluValThr 1640
 5201 ACACCTGGGGCCCCACACTCTCTCTGATACCGCTTGGGCTCTGTATTACCAACGAGGTG 5260
 1641 LeuThrHisProValThrLysTyrIleAlaThrCysMetGlnAlaAspLeuGluValMet 1660
 5261 CTCACGATCTCTGTGGAATAATACATCGGCACCTGATGCAAGCCGACCTTGGGTGATG 5320

Qy 1661 ThrSerThrTrpValLeuAlaGlyGlyValLeuAlaValAlaAlaTyrCysLeuAla 1680
Db 5321 ACCAGCAGCGGGCTTGTAGCTGGGGGGTCTTGGCGCGCTCGCGTACTCCCTGGCG 5380
Qy 1681 ThrGlyCysValCysIleIleGlyArgLeuHisIleAsnGlnArgAlaValAlaPro 1700
Db 5381 ACCGGGTGTTTGCATCATCGCGCGCTTGCAGTTAACCGAGCGCGCTGCTGACCG 5440
Qy 1701 AspGluGluValLeuTrpGluAlaPheAspGluMetGluGluCysAlaSerArgAlaAla 1720
Db 5441 GACAAGGAGGCTCTATGAGGCTTTTGTAGATGAGAGGAGGAGTGTCTCTAGAGCGCT 5500
Qy 1721 LeuIleGluGluGlyGlnArgIleAlaGluMetLeuIleCysSerIleGlnGlyLeuLeu 1740
Db 5501 CTCAATTCAAGAGGGGCGAGATAGCGAGATCTGAAGTCCAAGATCCAAGGCTTATTG 5560
Qy 1741 GlnGlnAlaSerIleGlnAlaGlnAspIleGlnProThrValGlnAlaSerTrpProIle 1760
Db 5561 CAGCAGCTTCCAAACAGCTCAAGCATCAACACCGCTGTGAGGCTTCTTGGGCCAAG 5620
Qy 1761 ValGluGlnPheTrpAlaIleHisMetTrpAsnPhelIleSerGlyIleGlnTrpLeuAla 1780
Db 5621 GTAGAGCAATTCCTGGGCCAAACACATGTGAATCTCATCAGCGGCATTCATACCTCGCA 5680
Qy 1781 GlyLeuSerThrLeuProGlyAsnProAlaValAlaSerMetMetAlaPheSerAlaAla 1800
Db 5681 GGACTATCAACACTGCGCAGGAAACCTGCTGTAGCTTCAATGATGCAATTCAGTCCGCC 5740
Qy 1801 LeuThrSerProLeuSerThrSerThrIleLeuLeuAsnIleLeuGlyTrpLeu 1820
Db 5741 CTCACGAGTCCGTTGCAACTAGCACACATATCTTCTCAACATTTTGGGGGCTGGCTA 5800
Qy 1821 AlaSerGlnIleAlaProProAlaGlyAlaThrGlyPheValValSerGlyLeuValGly 1840
Db 5801 GCATCCCAATTCGGCTCCCGCGGGGCTACCGGCTTCTGCTGAGTGGCTGTGGG 5860
Qy 1841 AlaAlaValGlySerIleGlyLeuGlyLeuValLeuValAspIleLeuAlaGlyTrpGly 1860
Db 5861 GCTGCGGTAGGCAGCATAGGCTTGGGTAAAGTCTGCTGTGACATCTCTGGCAGGATGGT 5920
Qy 1861 AlaGlyIleSerGlyAlaLeuValAlaPheLeuIleMetSerGlyGluIleProSerMet 1880
Db 5921 CGGGGATTTTCGGGGGCTCTGCTGCGCATTCAGATCATGTCTGGGAGAGCCCTCATG 5980
Qy 1881 GluAspValValAsnLeuLeuProGlyIleLeuSerProGlyAlaLeuValValGlyVal 1900
Db 5981 GAGATGTTGTCACCTGCTGCTGGAATCTCTCTCGGGTGCCTTGGTGGTGGAGTC 6040
Qy 1901 IleCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrpMet 1920
Db 6041 ATCTGCGCGCCATCTGCGCGCACACGCTGGGACCGGGGAGGCGCTGTCCAATGGATG 6100
Qy 1921 AsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaProThrHisTrpValThr 1940
Db 6101 AATAGGCTCATTCCTTGTCTCAGAGGAAACACGCTCGCCCCACCCACTACGTGAGC 6160
Qy 1941 GluSerAspAlaSerGlnArgValThrGlnLeuLeuGlySerLeuThrIleThrSerLeu 1960
Db 6161 GAGTGGATGCTGCGAGCGGTGACCAACTACTTGGCTCCCTTACCATAACAGCCCTG 6220
Qy 1961 LeuArgArgLeuHisAsnTrpIleThrGluAspCysProIleProCysGlyGlySerTrp 1980
Db 6221 CTCAGGAGACTCCACAACTGGATTACTGAAGACTGCCCATCTCCATGACGCGGCTCGTG 6280
Qy 1981 LeuArgAspValTrpAspTrpValCysThrIleLeuThrAspPheIleAsnTrpLeuThr 2000
Db 6281 CTCGGGATGTGGGATGGGTTTGCCATCTTAACAGACTTTAAAAAATCGGCTGACC 6340
Qy 2001 SerIleLeuPheProIleMetProGlyLeuProPheValSerCysGlnIleGlyTrpIle 2020
Db 6341 TCCAAATGTTCCAAAGATGCTGGTCTCCCTTTATCTTGTCAAAAGGGGTACAAG 6400
Qy 2021 GlyValTrpAlaGlyThrGlyIleMetThrThrArgCysProCysGlyAlaAsnIleSer 2040

Db 6401 GCGGTGTGGGCTGGCACTGGTATCATGACCAACGCTGTCTTGGCGGCCAATATCTCT 6460
Qy 2041 GlyAsnValArgLeuGlySerMetArgIleThrGlyProIleCysMetAsnIleTrp 2060
Db 6461 GGCATGTCCGCTGGGCTCCATGAGATTTAGCGGGCCCAAACTGATGATATCTGG 6520
Qy 2061 GlnGlyThrPheProIleAsnCysTrpThrGluGlyGlnCysValProIleProIlePro 2080
Db 6521 CAGGGAGCCTTTCCCATCAATTTGTACACGAGGGCCAGTGTGCGTGCAGAAACCGCACCA 6580
Qy 2081 AsnPheIleValAlaIleTrpArgValAlaAlaSerGluTrpAlaGluValThrGlnHis 2100
Db 6581 AACTTTAAGATCGCATCTCGAGGGTGGCGGCTCAGAGTACGCGAGGTGACGAGCAC 6640
Qy 2101 GlySerTrpHisTrpIleThrGlyLeuThrThrAspAsnLeuIleValProCysGlnLeu 2120
Db 6641 GGGTCATACCATACATACAGAGACTTACCATGTACTTGAAGTTCTTGGCCAACTA 6700
Qy 2121 ProSerProGluPhePheSerTrpValAspGlyValGlnIleHisArgPheAlaProThr 2140
Db 6701 CCTTCTCCAGAGTCTCTTTTCTGGGTGGAGCGAGTGCAGATCCATAGGTTTGGCCCAVA 6760
Qy 2141 ProIleProPhePheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal 2160
Db 6761 CCGAAGCCGCTTTTTCGGGATGAGGCTCTGCTTCTGCTTGGGCTTAATTCATTTGTCGTC 6820
Qy 2161 GlySerGlnLeuProCysAspProGluProAspThrAspValIleMetSerMetLeuThr 2180
Db 6821 GGGTCTCAGCTCTCTTTCGATCTTGAACCTGACACAGAGTATGTGACGTCCATGCTAAACA 6880
Qy 2181 AspProSerHisIleThrAlaGluThrAlaAlaArgArgLeuAlaArgGlySerProPro 2200
Db 6881 GACCCATCCCATATCACGCGGAGACTGCGAGCGGGCTTGGCAGCGGGTCAACCCCG 6940
Qy 2201 SerGluAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThr 2220
Db 6941 TCCGAGGCAAGCTCTCTCAGCGAGCCAGCTATCGGCACCATCGCTGCGAGCCACTGCACC 7000
Qy 2221 ThrHisGlyLeuAlaTrpAspValAspMetValAspAlaAsnLeuPheMetGlyGlyAsp 2240
Db 7001 ACCCAGCGGAGCGCTTATGATGTGACATGTGGATGCGCACTGTTCATGGGGGCGAT 7060
Qy 2241 ValThrArgIleGluSerGlySerIleValValValLeuAspSerLeuAspProMetVal 2260
Db 7061 GTGACCCGATAGAGTCTGAGTCCAAAGTGGTCTGTTCTGGAATCTCTCGACCCCAATGTC 7120
Qy 2261 GluGluArgSerAspLeuGluProSerIleProSerGluTrpMetLeuProIleIleArg 2280
Db 7121 GAAGAAAGGAGCGACTTTCAGCTTCGATACCATCGGAATATATGCTCCCAAGAGAGA 7180
Qy 2281 PheProProAlaLeuProAlaTrpAlaArgProAspTrpAsnProProLeuValGluSer 2300
Db 7181 TTCCACACGAGCTTACCGGCTTGGGCGCGCTGTATTAACCCACCCGCTTGTGGAAATCG 7240
Qy 2301 TrpIleArgProAspTrpGlnProAlaThrValAlaGlyCysAlaLeuProProArg 2320
Db 7241 TGAAGAGGCGAGATTAACACCGGCCACTGTGTGGGGCTGGCTCTCCCCCCCCCTAAG 7300
Qy 2321 LysThrProThrProProArgArgArgThrValGlyLeuSerGluAspSerIle 2340
Db 7301 AAAACCCGAGCGCTCCCAAGGAGAGCGCGGACAGTGGGTCTGAGTGAGAGCTCCATA 7360
Qy 2341 GlyAspAlaLeuGlnGlnLeuAlaIleIleYssPheGlyGlnProProProSerGlyAsp 2360
Db 7361 GCAGATGCCCTACAACAGCTGGCCATCAAGTCTCTTGGCCAGCCCCCCCCCAAGCGCGAT 7420
Qy 2361 SerGlyLeuSerThrGlyAlaGlyAlaAspSerGlySerGlnThrProProAspGlu 2380
Db 7421 TCAGGCTTTTCCACGGGGGCGAGCGAGCCGATTCGGCAGTTCGGAGCGCCCCCCCCATGAG 7480
Qy 2381 LeuAlaLeuSerGluThrGlySerIleSerSerMetProProLeuGluGlyLeuGly 2400

2761	AspLeuValValIleSerGluSerGlnGlyThrGluGluAspGluAArgAsnLeuArgAla	2780
8621	GACTTGGTTGTCATCTCAGAGAGCCAGGGGACCCAGGAGGACGACGCGAACTTGAGAGCC	8680
2781	PheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProArgProGluTyr	2800
8681	TTCAACGAGGCTATGACACAGGTATCTTGCCCTCTCTGGTGACCCCCCCAGACCCGGAATAT	8740
2801	AspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaLeuGlyProGlnGly	2820
8741	GACCTGAGCTGATAACATCTTGCTCTCAATGTGTCTGGGGGTTGGGCCCCCAAGGC	8800
2821	ArgArgArgTyrTyrLeuThrArgAspProThrThrProIleAlaAArgAlaAArgAlaAArgGlu	2840
8801	CGCCGACAGATACTACCTGACACAGAGACCTTACCACTCCAAATCGCCCGGGCTGCCCTGGGAA	8860
2841	ThrValArgHisSerProValAsnSerTrpLeuGlyAsnIleIleGlnTyrAlaProThr	2860
8861	ACAGTTAGACACTCCCTGTCTCAATTCATGGCTAGGAAACATCATCCAGTACGCCCAACC	8920
2861	IleTrpAlaArgMetValLeuMetThrHisPheSerIleLeuMetAlaGlnAspThr	2880
8921	ATATGGGCTCGCATGCTCTCATGACACACTTCTTCTCAATCTCATGGCCCAAGATACT	8980
2881	LeuAspGlnAsnLeuAsnPheGluMetTyrGlyAlaValTyrSerValSerProLeuAsp	2900
8981	CTGACACAGAACCTTCAACTTTGAGATGTACGAGCGGTGTACTCGTGAGTCTCTTGAC	9040
2901	LeuProAlaIleIleGluArgLeuHisGlyLeuAspAlaPheSerLeuHisThrTyrThr	2920
9041	CTCCAGCCATAATTGAAGGTTACACGGGCTTGACGGCTTCTCTCTGCACACATACACT	9100
2921	ProHisGluLeuThrArgValAlaSerAlaLeuArgLysLeuGlyAlaProProLeuArg	2940
9101	CCCCACCAACTTGACACGGGTGGCTTACGCCCTTCAGAAACTTGGGGCGCCACCCCTCAGA	9160
2941	AlaTrpLysSerArgAlaAArgAlaValArgAlaSerLeuIleSerArgGlyGlyArgAla	2960
9161	CGTGGAAGACCCGGGCACGTGCAGTCAGGGGGTCCCTCATCTCCGTGGGGGAGAGCG	9220
2961	AlaValCysGlyArgTyrLeuPheAsnTrpAlaValLysThrLysLeuLysLeuThrPro	2980
9221	GCCGTTTTGCGGCCGATATCTCTTCAACTGGGCGGTGAAGACCAAGCTCAAACTCATCTCCA	9280
2981	LeuProGluAlaArgLeuLeuAspLeuSerSerTrpPheThrValGlyAlaGlyGlyGly	3000
9281	TTGCGGAAGCGCGCTCTCGATTTATCCAGCTGGTTTACTGTCTGGCGCCGGCGGGGCG	9340
3001	AspIleTyrHisSerValSerArgAlaArgProArgLeuLeuLeuPheGlyLeuLeuLeu	3020
9341	GACATTATACACGGCTGTGCGGTGCCCGACCCCGCTTATTACTTCTTGGGCTACTCTCTA	9400
3021	LeuPheValGlyValGlyLeuPheLeuLeuProAlaArg	3033
9401	CTTTTGTAGGGGTAGGCTTTTCTCTACTCCCGCTCG	9439

RESULT 3
UIS-07-925-695-7

US-07-925-695-7
: Sequence 7. Application US/07925695

Sequence /, Application
Patent No. 5428145

Patent No. 5428143

APPLICANT: OKAMOTO, Hiroaki

APPLICANT: OKAMOTO, HIROAKI
APPLICANT: NAKAMURA, Tetsuo

APPLICANTI: NAKAMURA, Tetsuo
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOME,

NON-A, NON-B HEPATITIS: POLYNUCLEOTIDES, H

NO.	TITLE OF INVENTION:	POIN
1	TITLE OF INVENTION:	DET
2	TITLE OF INVENTION:	DET

;; TITLE OF INVENTION: DE
;; NUMBER OF SEQUENCES: 9

; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Beveridge, DeGrandi, Weilacher & Young

STREET: 1850 M Street, N.W., Suite 800

SIREEL: 1830 M SC
CITY: Washington

CITY: WASHINGTON
 STATE: D.C.

Copied from 09540843 on 05/19/2004

COUNTRY: US
 ZIP: 20036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/925,695
 FILING DATE: 19920807
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 287402/91
 FILING DATE: 09-AUG-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 360441/91
 FILING DATE: 05-DEC-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: weilacher, Robert G.
 REGISTRATION NUMBER: 20,531
 REFERENCE/DOCKET NUMBER: 06/87-48009
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 659-2811
 TELEFAX: (202) 659-1462
 TELEX: WUI 64470
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 9511 base pairs
 TYPE: NUCLEIC ACID
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-07-925-695-7

Alignment Scores:

Pred. No.: 0 Length: 9511
 Score: 13810.00 Matches: 2523
 Percent Similarity: 90.93% Conservative: 235
 Best Local Similarity: 83.18% Mismatches: 275
 Query Match: 85.67% Indels: 0
 DB: 1 Gaps: 0

US-09-980-559-2 (1-3033) x US-07-925-695-7 (1-9511)

QY 1 MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGln 20
 DB 342 ATGAGCACAAATCTTAAACCTCAAGGAAACCAAGAAACCAACACCCGCCGCCACAG 401
 QY 21 AspValLysPheProGlyGlyGlyGlnIleValGlyGlyValTyrLysLeuProArg 40
 DB 402 GACGTTAAAGTTCCCGGCTGGCGGTGACATCGTTGGCGGAGTTTACTTGTCTGCCGCGCAGG 461
 QY 41 GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly 60
 DB 462 GGGCCCGAGGTGGGTGTGCGCGGCGACAGGAAGACTTCYAGCGGATCCCGCGCGTGA 521
 QY 61 ArgArgGlnProIleProLysAspArgArgSerThrGlyLysSerTyrGlyLysProGly 80
 DB 522 CGAGCCGAGCCCATCCCGAAGATCGCGCTCCACCGGCAAGTCTCTGGGGAAGCCAGGA 581
 QY 81 TyrProTyrProLeuTyrGlyAsnGluGlyLeuGlyTyrAlaGlyTyrLeuLeuSerPro 100
 DB 582 TATCCTTGGCCCTGTACGGAACAGAGGGTTGCGGCTGGCGGTGGGTCTCTGTCTGCC 641
 QY 101 ArgGlySerArgProSerTyrProAsnAspProArgHisArgSerArgAsnValGly 120
 DB 642 CGCGGGTCTCGTCTTCTTGGGGCCCCACCGACCCCGCGCATAGATCAGCGCAATTTGGG 701
 QY 121 LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProValVal 140
 DB 702 AGAGTCATCATACCAATACGTGTGTTTTCGACCTCATGGGGTACATCCCTGTCTG 761
 QY 141 GlyAlaProLeuGlyGlyValAlaAlaArgAlaLeuAlaHisGlyValArgValLeuGluAsp 160

DB 762 GCGCGCCCGGTGGAGCGCTCGCAGAGCTCTGGCACACGGTGTATTAGGCTCTCGGAGGAC 821
 QY 161 GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla 180
 DB 822 GGATAAATACGCAACAGGAATTTACC CGGTGCTCTTTTCTATCTTTTGTCTGCT 881
 QY 181 LeuLeuSerCysIleThrThrProValSerAlaAlaGluValLysAsnIleSerThrGly 200
 DB 882 CTTCTGTCTGTCTGTCACARTGCCAGTGTCTGCAGTGGAGTCAAGAACATYAGTCTTCTAGC 941
 QY 201 TyrMetValThrAsnAspCysThrAsnAspSerIleThrTyrGlnLeuGlnAlaAlaVal 220
 DB 942 TACTACGCCCACTAATGATTCCTCAACCAACAGCATCCTCGGAGCTCACTGACGCGAGTT 1001
 QY 221 LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTyrIle 240
 DB 1002 CTCCATCTTCTGGATGCGTCCATGTGAGAAATAYGGCACCTTGTCTGCTGGATA 1061
 QY 241 ProValSerProAsnValAlaValGlnArgProGlyAlaLeuThrGlnGlyLeuArgThr 260
 DB 1062 CAAAGTAAACACCCACCGTGTGTAACACCGCGGTGCGTCACTCGTAGCTGCGAACA 1121
 QY 261 HisIleAspMetValValMetSerAlaThrLeuCysSerAlaLeuTyrValGlyAspLeu 280
 DB 1122 CACGTCACATGATCGTAATGGCAGCTACGGCTGCTCGGCTTGTATGTGGAGATGTG 1181
 QY 281 CysGlyGlyValMetLeuAlaAlaGlnMetPheIleValSerProGlnHisHisTyrPhe 300
 DB 1182 TCGGGCGCGTGATGATCTATCGCAGGCTTTCATGTGTATCACCACACGCCAACACTTC 1241
 QY 301 ValGlnAspCysAsnCysSerIleTyrProGlyThrIleThrGlyHisArgMetAlaTyr 320
 DB 1242 ACCCAAGAGTGAACCTGTTCCATCTACCAAGTGCATCACCGGCCATCGATGGCATGG 1301
 QY 321 AspMetMetMetAsnTyrSerProThrAlaThrMetIleLeuAlaTyrAlaMetArgVal 340
 DB 1302 GACATGATCTTACCTGGTCTCCCACTCTTTCATGATCTCGCTTCTGCTGCGGGA 1361
 QY 341 ProGluValIleIleAspIleIleSerGlyAlaHisTyrGlyValMetPheGlyLeuAla 360
 DB 1362 CCGGACTGCTCTCGAATATYTTTCGGCGGCGCATTTGGGTGTGGGTGTGGCTTGGGC 1421
 QY 361 TyrPheSerMetGlnGlyAlaTyrAlaLysValValIleLeuLeuLeuAlaAlaGly 380
 DB 1422 TATTTCTCCATCGAGGAGCGTGGGCAAGTCRTYGCCATCTCTCTCTTGTGGGGA 1481
 QY 381 ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr 400
 DB 1482 GTGGATGCACCCCTATTTCCACGCGYACGAGCGGTCTGCCGTCYCMKGGGWTCKCT 1541
 QY 401 SerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTyr 420
 DB 1542 RGCTCTTTTAMTACTGGTCCCAAGCAACCTCYATTTRATCAACACCAATGGCAGCTGG 1601
 QY 421 HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer 440
 DB 1602 CACATAAACCGGACTGCCCTCAATTCATATGACAGCYTASAGAGCGGTTTCTGCTGCC 1661
 QY 441 LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg 460
 DB 1662 YTGKTTTACMWCRCARGTTCAACAGCTCTGGCTGCCCGGAGCGCTTGTCTCTGCCCG 1721
 QY 461 SerIleGluAlaPheArgValGlyTyrGlyAlaLeuGlnTyrGluAspAsnValThrAsn 480
 DB 1722 GGGTGGAGAVTTTGGCATCGGTGGGGAACCTTGGAAATACAAACACCAAGCTCACAAC 1781
 QY 481 ProGluAspMetArgProTyrCysTyrHisTyrProProArgGlnCysGlyValValSer 500
 DB 1782 GATRGACATGAGCGCGTACTGCTGGCATTACCCCGGAGCGCTTGGCGCATCTGCTCCG 1841
 QY 501 AlaLysThrValCysGlyProValTyrCysPheThrProSerProValValValGlyThr 520
 DB 1842 GCTAGGACGGTTGGCGACCGGTCTATTGTGTTTCAACCCCTAGCCCTGCTGCTGGGCACC 1901

QY	521	ThrAspArgLeuGlyAlaProThrTyrThrTyrGluAsnGluThrAspValPheLeu	540
DB	1902	ACTGACAGCAGGCGCTACCCACCTAACCTGGGGRGAAACGAGACCGATGCTCTCCG	1961
QY	541	LeuAsnSerThrArgProProLeuGlySerTrpPheGlyCysThrTrpMetAsnSerSer	560
DB	1962	CTRAATAGCACAGACCCCGCGAGGAGCTTGGTTCGGCTGCACATGGATGAACGGGACT	2021
QY	561	GlyTyrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnLaser	580
DB	2022	GGGTTCACTAAGACATCGCGTGCCACCACTTCCGCAATAGGAAGACTACACAGCACT	2081
QY	581	ThrAspLeuLeuCysProThrAspCysPheArgLysHisProAspThrThrTyrLeuLys	600
DB	2082	CTCGATTATTTGGCCCCACAGACTGTTTATAGAGCCACCCAGATGCTATCTATTAAG	2141
QY	601	CysGlySerGlyProTrpLeuThrProArgCysLeuLeuAspTyrProTyrArgLeuTrp	620
DB	2142	TGTGGAGCAGGCGCTTGGTTAACTCCACAGTGCTTGGTAGACTACCTTATAGRYTGG	2201
QY	621	HisTyrProCysThrValAsnTyrThrIlePheLysIleArgMetTyrValGlyGlyVal	640
DB	2202	CATTATCCGTGCACTGTAACTTCACTTAAAGCGCGGATGTATGTAGGAGGGTG	2261
QY	641	GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp	660
DB	2262	GAGCATCGATTCCTCCGAGCATGCACTTCCACGCGGAGATCGCTGCAGACTGGAAGAT	2321
QY	661	ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGluTrpAlaIleLeuPro	680
DB	2322	AGGATAGGGGCGAGCAGACTCCACTGCTGCACTTCCACTACTGAGTGGGCGGTCCCA	2381
QY	681	CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisIleLeuHisGlnAsnIle	700
DB	2382	TGCTCTCTCTGACCTACCCAGCACTATCCACTTGGCTTATGCACTCCACCAAAACATC	2441
QY	701	ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrLysTyrIleValArgTrp	720
DB	2442	GTGGACGTGAGTACCTTACGAGCTTCTCCGGCTCTGCAAGAGTACATCTGGAAGTG	2501
QY	721	GluTrpValIleLeuLeuPheLeuLeuAlaAspAlaArgValCysAlaCysLeuTrp	740
DB	2502	GAGTGGGTGATCCTCTTCTTGTGTCGAGAGCCAGGTCGTGTCATGCCCTTTGG	2561
QY	741	MetLeuIleLeuLeuGlyGlnAlaGluAlaLeuGlnLysLeuValIleLeuHisAla	760
DB	2562	ATGCTCAWCATACCTGGGCCAAGCGCAAGCGCGCTTGAGAAGCTCATCTTTGCACCTC	2621
QY	761	AlaSerAlaAlaSerCysAsnGlyPheLeuTyrPheValIlePhePheValAlaAlaTrp	780
DB	2622	GCTAGYGTCTGTAGTGCATAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2681
QY	781	TyrIleLysGlyArgValValProLeuAlaThrTyrSerLeuThrGlyLeuTrpSerPhe	800
DB	2682	TACTTAAAGGCGAGGTGGTCCCGTGGCCAGCTACTCTGCTBCTCGGCTTGTGCTCTC	2741
QY	801	SerLeuLeuLeuAlaLeuProGlnGlnAlaTyrAlaTyrAspAlaSerValHisGly	820
DB	2742	CTCTCTCTAGTCTGCTGCTTACCACAGCAGCTTATGCTTGGACGCTGCTGAACAGGG	2801
QY	821	GlnIleGlyAlaAlaLeuLeuValMetIleThrLeuPheThrLeuThrProGlyTyrLys	840
DB	2802	GACTGGGGTGGCCATATTAGTAAATATTATCACTTTACTCTTACCCAGCATACAG	2861
QY	841	ThrLeuLeuSerArgPheLeuTrpTrpLeuCysTyrLeuLeuLeuThrLeuGlyGluAlaMet	860
DB	2862	ATCCTCTGAGCGCTTCACTGTGGTGTGCTGCTTACATGCTGCTTGGCGGAGGCCAG	2921
QY	861	ValGlnGluTrpAlaProProMetGlnValArgGlyGlyArgAspGlyIleIleTrpAla	880
DB	2922	ATTGAGCAATGGGTTCCTCCCTCGAGGTCCGAGGGGGGTGTCAGGATCATCTGGGTG	2981
QY	881	ValAlaIlePheTyrProGlyValValPheAspIleThrLysThrLeuLeuAlaValLeu	900
DB	2982	GCTGTGCTTCTACCCAGCCTTGTGTTTGGTTCACGAAATGTTGTAGCAATCTCTG	3041
QY	901	GlyProAlaTyrLeuLeuLysGlyAlaLeuThrArgValProTyrPheValArgAlaHis	920
DB	3042	GGGCTGCTCTCTCTTARAGCGTCTCTGTCACGGATACCGTACTTGTGAGGGGCCAC	3101
QY	921	AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyGlyArgTyrValGlnMet	940
DB	3102	GCTTGTCTACGAGTGTGTACCTGTGTAACCTCGCGGGGCTAGGTACATCCAGATG	3161
QY	941	AlaLeuLeuAlaLeuGlyArgTyrThrGlyTyrIleTyrAspHisLeuThrProMet	960
DB	3162	CTGTTRATCACCATAGGCGAGATGGACCGCACTTACATCTACGACCACTCTCCCTTTA	3221
QY	961	SerAspTrpAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIleIlePhe	980
DB	3222	TCAACTTGGGCGGCCAGGTTTTCGGGACCTTGGCAATCGCGTGGAGCTGTGGTGTTC	3281
QY	981	SerProMetGluLysLysValIleValTrpGlyAlaGluThrAlaAlaCysGlyAspIle	1000
DB	3282	AGCCCAATCGAGAAGATGCTTGTGGGGGCTGACACAGTGGCGTGTGAGACATC	3341
QY	1001	LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp	1020
DB	3342	CTGCATGGCTCTCCGCTCTCCGAGGCTAGGTAGGARGTCTTCTCGCGCCCTGCCGAC	3401
QY	1021	GlyTyrThrSerLysGlyTrpSerLeuAlaProIleThrAlaTyrAlaGlnGlnThr	1040
DB	3402	GGCTACACCTCCAAAGGGTGGAAKCTCTTACGTCTTACCTTACCTACCTACCAACT	3461
QY	1041	ArgGlyLeuLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGluGlnAla	1060
DB	3462	CGTGGTCTCTCGGTGCTATCTGTGTGCTTACGCGCGCGGACCAAAATGAGCAGCT	3521
QY	1061	GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerLysGly	1080
DB	3522	GGGAGGTCCAGGCTCTGCTCTCCGTCCACAAACTTCTTGGGGACATCCATTTCCGGC	3581
QY	1081	ValLeuTrpThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro	1100
DB	3582	GTCTCTTGACAGTATATACCGGGCTGTGTAAAGACCTTGGCGCGCCCAAGGAGCA	3641
QY	1101	ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTyrProSerProGly	1120
DB	3642	GTCTCTTGACAGTATATACCGGGCTGTGTAAAGACCTTGGCGCGCCCAAGGAGCA	3701
QY	1121	ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArgAsn	1140
DB	3702	ACTAAGTCTATGGACCCCTGTACCTCGGGGCGGTAGACCTCTACCTGGTCAACCAAC	3761
QY	1141	AlaAspValIleProAlaArgArgGlyAspLysArgGlyAlaLeuLeuSerProArg	1160
DB	3762	GCTGATGTCTTCCGCTCGGAGAAAGATGACCGAGGGGGTGTAGTCTTCTCGCAAGG	3821
QY	1161	ProLeuSerThrLeuLysGlySerSerGlyProValLeuCysProArgGlyHisAla	1180
DB	3822	CCCTCTCAACCTCAAGGATCATCCGAGGGGCGGTGTCTGCTCTGAGGGGACACGCC	3881
QY	1181	ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle	1200
DB	3882	GTGGGCTTGTTCAGAGCGCGGTGTGTCGAGGGGTGTAGCCAAATCTATTGACTTATC	3941
QY	1201	ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro	1220
DB	3942	CCGCTGCAATCACTCGCATCTCGCCACGAGCGCCAGGTTCTCTGACAACTGTCGCCG	4001
QY	1221	ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLys	1240
DB	4002	CCAGCTGTGCGCCGAGTCTTACAGGTGGGTACTTGCACCCACCAACAGCAGCGGGAAG	4061
QY	1241	SerThrLysValProValAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnPro	1260

5142	GCAAAGGCCCTCCTCTTCTCGGACGTGATGGGAAGTGCTAACTAGCGCTCAACCT	5201
1621	ThrLeuValGlyProThrProLeuLeuTyArgLeuGlySerValThrAsnGluValThr	1640
5202	ACACTGACTGGTCCACCCTCTCTGTACCGCTTGGGTGCCTGACCAATGAGGTYACC	5261
1641	LeuThrHisProValThrLysTyrlleAlaThrCysMetGlnAlaAaspLeuGluValMet	1660
5262	TTCAGCACCCCGGTGACAAAATACATGCCACGTGTCATGCAAGCTGACCTYGAGATCATG	5321
1661	ThrSerThrTrpValLeuAlaGlyVglValLeuAlaValAlaAlaTyrcysLeuAla	1680
5322	ACAAGCTCATGGGTCTCGCGGGGGGGTGTCTAGCCCGCTGGCAGCTTTACTGCCTGGCG	5381
1681	ThrGlyCysValCysIleileGlyArgLeuHisIleAsnGlnArgAlaValValAlaPro	1700
5382	ACTGGCTGCATTCCATCATTTGGCGGCTACACCTGAATGATCGGGTGGTTGTGRCCCY	5441
1701	AspLysGluValLeuTyrgluAlaPheAspGluMetGluGluCysAlaSerArghAlaLa	1720
5442	GACAAGGARATCTTTATATGAGCGCTTTGATGAGATGGAAGAATGCCCTCCAAGACCGCC	5501
1721	LeuileGluGluGlnArgIlealacLumetLeuLysSerLysIleleGlnCylLeuLeu	1740
5502	CTCATTGGGAAGGGCAGCGATGCGGAGATGCTCAAAATCTTAAGATACAAGGCTTCCTTA	5561
1741	GlnGlnAlaSerLysGlnAlaGlnAspileGlnProThrValGlnAlaSerTrpProLys	1760
5562	CAACAGGCCACAGGCAAGCTCAAGRCATRCAGCCAGCTATACAGTCATCATGGCCCCAAG	5621
1761	ValGluGlnPheTrpAlaLysHieMetTrpAsnPheIleSerGlyIleGlnTyrlleuAla	1780
5622	CTTGAACAATTTTGGGCCAAACACATGTGGAACCTTCATCAGTGGTATACAGTACCTTAGCA	5681
1781	GlyLeuSerThrLeuProGlyAsnProAlaValAlaSerMetMetAlaPheSerAlaLa	1800
5682	GGACTCTCACCCCTACCGGGAAATCTCGCATRGTCATCAATGATGGCTTTTAGCCCGCGG	5741
1801	LeuThrSerProLeuSerThrSerThrThrIleLeuLeuAsnIleLeuGlyClyTrpLeu	1820
5742	CTGACTAGCCCATACCCACCAGCACCATCTCTTTGAACATCATGGGAGGATGCTTG	5801
1821	AlaSerGlnIleAlaProProAlaGlyAlaThrGlyPheValValSerGlyLeuValGly	1840
5802	GCCTCYCAGATTGCCCCCTCGCGAGCCACACYGGCTTCGTGTGTCAGTGGTCTAGTGGGG	5861
1841	AlaAlaValGlySerIleGlyLeuGlyLysValLeuValAspIleLeuAlaGlyTyrgLy	1860
5862	CGCGCGGTGGAAGCATAGGCTCGGTGAAGATACCTGGTGGACGTTTTCGCGGGGTACGCG	5921
1861	AlaGlyIleSerGlyAlaLeuValAlaPheLysIleMetSerGlyGluLysProSerMet	1880
5922	CAGGCAATTCAGGGGCCCTCGTAGCTTTTAAGATCATGAGCGCGGAGAACCCACGGTA	5981
1881	GluAspValValAsnLeuLeuProGlyIleLeuSerProGlyAlaLeuValValGlyVal	1900
5982	GAAGACGTGTGGAATCTCTCGCTGCTATYCTGTCTCTGTGGCGTGTGATGGGAGTGC	6041
1901	IleCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrpMet	1920
6042	ATCTGTGAGCAATYCTCGCGCCGACATCGCTCGGTCAAGGAGAGGRCGGTCCAGTGGGATG	6101
1921	AsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaProThrHisTyrlvalThr	1940
6102	AACGACTGATCGCTTCGCTTCAGGGGGAACCAACGTTTGGCCCTTACCACCTACGTGGTG	6161
1941	GluSerAspAlaSerGlnArgValThrglnLeuLeuGlySerLeuThrIleThrSerLeu	1960
6162	GAGCTCGAGCTTCACAGCGGTTRAGCCAGGTGCTGAGTTCACTTACAATTTACCAGCTTA	6221
1961	LeuArgArgLeuHisAsnTrpIleThrglnAspCysProIleProCysGlyClysSerTrp	1980
6222	CTTAGGAGACTACATGCTCGATCACTGAAGTTGCCARTCCATGCTCGGGGTCTTGG	6281

QY	1981	LeuArgAspValTyrAspThrValCysThrIleLeuThrAspPheLysThrLeuThr	2000	2341	GlyAspAlaLeuGlnGlnLeuAlaIleLysSerPheGlyGlnProProProSerGlyAsp	2360
DB	6282	CTCCAGGACATTTGGGATTTGGTTTCTCATCTCCACAGACTTAAACACTGGCTGTCT	6341	7362	GAGGGGRTCTCTCAGGAGATGGCTGACAAAGTCTCAGCCCTCTCAACACCAATGAC	7421
QY	2001	SerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTyrLys	2020	2361	SerGlyLeuSerThrGlyAlaAlaAspSerGlySerGlnThrProProAspGlu	2380
DB	6342	TCNAATTAATCTCCCAAGATGCCCGCATCTCTTATCTCTGCGCAGAGGATACAG	6401	7422	TCGGGTCACTCCACTGGAGCGATACCGGAGGAGACATCGTCCAGCAACCTCTGACGAG	7481
QY	2021	GlyValTyrAlaGlyThrGlyIleMetThrThrArgCysProCysGlyAlaIleSer	2040	2381	LeuAlaLeuSerGluThrGlySerIleSerMetProProLeuGlnGlyLeuGly	2400
DB	6402	GGTGTATGGCTGCTACCGGTGTATGATGATGATGATGATGATGATGATGATGAT	6461	7482	ACTGGCGTCTCAGAACGGGGTCACTGCTCTCATGCTCCCTCTGAGGAGGAGCGGGA	7541
QY	2041	GlyAsnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTyr	2060	2401	AspProAspLeuGluProGlnValGluProGlnProProGlnGlyValAla	2420
DB	6462	GGCCATGTCCGATGGGACCATGAATAAAGGCGCCGAAAGACTTGTCTGAACCTGTGG	6521	7542	GACCCYGACCTGGAGTTGAACAGTGGATCGGCTCCCTCTCAGGGGGAGTGTGAG	7601
QY	2061	GlnGlyThrPheProIleAsnCysTyrThrGluGlyGlnCysValProLysProAlaPro	2080	2421	AlaProGlySerAspSerGlySerThrSerThrCysSerGluGluAspAspSerVal	2440
DB	6522	CAGGGACCTTCCCATTAATTTGTACAGAGAGGCCCTGCGTGCNAACCCCTCTCT	6581	7602	GTCATTGATTGGAGTCTTAAGTCGTGTCCACAGTCTCTGATCAAGAGGATTTCTGTATC	7661
QY	2081	AsnPheLysValAlaIleTyrArgValAlaAlaSerGluTyrAlaGluValThrGlnHis	2100	2441	CysCysSerMetSerTyrSerThrThrGlyAlaLeuIleThrProCysSerProGluGlu	2460
DB	6582	AATTACAAAGACCCCAATTTGGAGGGTGGCAGCGTGGAGTACGTTGAGGTTCACACAGCAT	6641	7662	TGCTGCTCTATGTCTATCTCTGAGCGGGGGCTCTATACACCATGTGGGCCGAGAG	7721
QY	2101	GlySerTyrHisTyrIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeu	2120	2461	GluLysLeuProIleAsnProLeuSerAsnSerLeuLeuArgTyrHisAsnLysValTyr	2480
DB	6642	GGCTCTTTCGCTATGTAACGGGTAAACAGTGAACACCTTAAGGTCTCTGCCAGGTA	6701	7722	GAGAGTTACCGATCAACCTCTGAGTAATTCGCTCATCGGTTCATTAAGGTGTAC	7781
QY	2121	ProSerProGluPhePheSerTyrValAspGlyValGlnIleHisArgPheAlaProThr	2140	2481	CysThrThrThrLysSerAlaSerLeuArgAlaLysLysValThrPheAspArgMetGln	2500
DB	6702	CCAGCTCCAGAAATTTCTCTGGGTGGAGGGGTGCNAATCCACCGATTTCGCCCGGTW	6761	7782	TCCACAACTCGAGGAGTCCCTCTCAGGGGCAAGAGAGTTCATTTTACAGGGTGCAG	7841
QY	2141	ProLysProPhePheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal	2160	2501	ValLeuAspSerTyrTyrAspSerValLeuLysAspIleLysLeuAlaLysLysVal	2520
DB	6762	CCAGGTCCCTTCTTTCGGGATGAGTAAACGTTCAACGTTAGGCTTAACTCTCTGCTC	6821	7842	GTGCTGGAGCGCATATGACTCATGCTCTGAGGAGGAGTAAAGCGGCCCTCTAAGGT	7901
QY	2161	GlySerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThr	2180	2521	ThrAlaArgLeuLeuThrMetGluGluAlaCysGlnLeuThrProHisSerAlaArg	2540
DB	6822	GGCTCTCAGCTCCCTTTCGATCTGAGCGCGGACCCGAGTACTGGCCTCTATGTTGACA	6881	7902	RGTCGAGGCTCTCTACAGTAGAGGAGGCTGCGGCTGACCCCGCCCTCTCCGCCAA	7961
QY	2181	AspProSerHisIleThrAlaGluThrAlaAlaArgArgLeuAlaArgGlySerProPro	2200	2541	SerLysTyrGlyPheGlyAlaLysGluValArgSerLeuSerGlyArgAlaValAsnHis	2560
DB	6882	GACCGTCCACATCACCGCKGAGGGCGGACCCAGCGGATTCGCAAGGGGATCTCCCCY	6941	7962	TCGCGATACGATTTGGGCAAGAGGTCGCGAGCTTATCCAGGAGGGCGGTAAACCAC	8021
QY	2201	SerGluAlaSerSerAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThr	2220	2561	IleLysSerValTyrLysAspLeuGluAspSerGluThrProIleProThrThrIle	2580
DB	6942	TCACAGGCTAGCTCCTCAGCGAGCCAGCTCTCTGCCCGCTCTTGAAGGCTACTGTACC	7001	8022	ATCCCGTCCGTGGGAGGACCTCTCGAAGACCAACRTACCCCAATTCACACATATC	8081
QY	2221	ThrHisGlyLysAlaTyrAspValAspMetValAspAlaAsnLeuPheMetGlyGlyAsp	2240	2581	MetAlaLysAsnGluValPheCysValAspProThrLysGlyLysLysAlaAlaArg	2600
DB	7002	ACCCATAAGACAGCATATGATGTGACATGGTGATGATGATGATGATGATGATGAT	7061	8082	ATGGCTTAAATGAGGTGTTCTGCTGATCCAACTAAGGTGGGAAAGCCAGCTCGC	8141
QY	2241	ValThrArgIleGluSerGlySerLysValValLeuAspSerLeuAspProMetVal	2260	2601	LeuIleValTyrProAspLeuGlyValArgValCysGluLysMetAlaLeuTyrAspIle	2620
DB	7062	GTGACCGGATTCAGTCTGACTCTAAGTGTATGCTTCTAGCTCCCTCGATTCATGACT	7121	8142	CTCATCGTATACCCGACCTTGGGTGAGGAGGATGCGGAGGATGCGCTCTATGATC	8201
QY	2261	GluGluArgSerAspLeuGluProSerIleProSerGluTyrMetLeuProLysLysArg	2280	2621	ThrGlnLysLeuProGlnAlaValMetGlyAlaSerTyrGlyPheGlnTyrSerProAla	2640
DB	7122	GAGGTAGAGGATGATCTGAGCGCTTCTGTACCATCAGAGTACCTGTATCAAGAGGAAAG	7181	8202	RCACAAAGGCTCCCAAGCGATATGGGGCCATCTATGGGTTCATTAATCTCTCCGCA	8261
QY	2281	PheProAlaLeuProAlaTyrAlaArgProAspTyrAsnProProLeuValGluSer	2300	2641	GlnArgValGluPheLeuLeuLysAlaTyrAlaGluLysLysAspProMetGlyPheSer	2660
DB	7182	TTCCACCGGCGTCTCTCTGGGCGGCTCCAGACTACCAATCTCTTTTGTATGAGAGCA	7241	8262	GAACGGGTGATTTCTCTCAAGCTTGGGAGGATGAGAGGACCAATGGGGTCTCG	8321
QY	2301	TyrLysArgProAspTyrGlnProAlaThrValAlaGlyCysAlaLeuProProArg	2320	2661	TyrAspThrArgCysPheAspSerThrValThrGluArgAspIleArgThrGluGluSer	2680
DB	7242	TGGNAGGCGGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT	7301	8322	TATGACACCGCTGCTTCTGACTCAACCGTCAAGGAGGAGGACATAAGAGAGGATTC	8381
QY	2321	LysThrProThrProProArgArgArgThrValGlyLeuSerGluAspSerIle	2340	2681	IleTyrArgAlaCysSerLeuProGluGluAlaHisThrAlaIleHisSerLeuThrGlu	2700
DB	7302	CAAAAGCAGTCTCCACCTCGAGCGCGCGGCAAAATCTCTGACCCAGGACRATGTG	7361	8382	ATATATCAGGCTGTTCTCTGCTCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	8441
				2701	ArgLeuTyrValGlyGlyProMetPheAsnSerLysGlyGlnThrCysGlyTyrArgArg	2720

```
Db 8442 AGACTTTACGTAGGAGGCCCATCACAAACAGCAAAAGGCAATCTCGCGCTACAGCGGT 8501
Qy 2721 CybArgAlaSerGlyValLeuThrThrSerMetGlyAsnThrThrCysTyrValLys 2740
Db 8502 TGCCGGCAGCGGKGTTCACACACAGATGGGAATACCAATGATGTTACATCAA 8561
Qy 2741 AlaLeuAlaAlaCysLysAlaAlaGlyIleAlaProThrMetLeuValCysGlyAsp 2760
Db 8562 GCCCTTGACAGGTGAAGGCTGCRGGATCGTGACCCCTGTTATGTTGTTGTTGGAGAC 8621
Qy 2761 AspLeuValValLysSerGlySerGlnGlyThrGluGluAspGluArgAla 2780
Db 8622 GACCTGGTCTCATCTCAGAGAGCAAGGTAAACGAGGAGGAGGAGGAGGAGGAGGAGGAG 8681
Qy 2781 PheThrGluAlaMetThrArgTyrSerAlaProGlyAspProProArgProGluTyr 2800
Db 8682 TTCAGGAGGCTATGACAGGATTCGCCCTCCCGGTGACCTCCAGACCGGAATAT 8741
Qy 2801 AspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaLeuGlyProGlnGly 2820
Db 8742 GACTTGGAGCTTAAATCTCTCTCTCAACGATATCGGTAGCGCTGGACTCTCGGGT 8801
Qy 2821 ArgArgArgTyrTyrLeuThrArgAspProThrThrProIleAlaAlaAlaTyrGlu 2840
Db 8802 CGCCGCGGTACTTCTTAACAGAGACCTTACCCTCCAAATCAACCGAGCTGCTGGAA 8861
Qy 2841 ThrValArgHisSerProValAlaSerTyrLeuGlyAsnIleIleGlnTyrAlaProThr 2860
Db 8862 ACAGTAAGACACTCCCTCTGTCATTTCTGGCTGGGCAACATCATCAGTACGCCCCACA 8921
Qy 2861 IleTyrAlaArgMetValLeuMetThrHisPheSerIleLeuMetAlaGlnAspThr 2880
Db 8922 ATCTGGGTCCGATGTCATATGATGACTCATCTTCTTCCATATATGGCCAGGACACT 8981
Qy 2881 LeuAspGlnAsnLeuAsnPheGluMetTyrGlyAlaValTyrSerValSerProLeuAsp 2900
Db 8982 CTGAACCAAAATCTCAATTTGAGATGTACGGGCGAGTATCTCGGTCAATCCATTAGAC 9041
Qy 2901 LeuProAlaIleIleGluArgLeuHisGlyLeuAspAlaPheSerLeuHisThrTyrThr 2920
Db 9042 CTACCGGCCATAATTGAAGGCTACATGGGCTTGAAGCCCTTTTCACTGCACACATACTCT 9101
Qy 2921 ProHisGluLeuThrArgValAlaSerAlaLeuArgLysLeuGlyAlaProProLeuArg 2940
Db 9102 CCCCAGCACTCTCAGGGTGGGAGCACTCTCAGAAACTTGGAGGGCTCCCTTAGA 9161
Qy 2941 AlaTyrLysSerArgAlaArgAlaValAlaSerLeuIleSerArgGlyGlyArgAla 2960
Db 9162 GCGTGAAGAGTCCGGCGGCTGCGTGAGAGCTTCACTCATGCCCAAGGAGCGGGCG 9221
Qy 2961 AlaValCysGlyArgTyrLeuPheAsnTyrAlaValLysThrLysLeuLysLeuThrPro 2980
Db 9222 GCATTTGTGGCGCTTACTCTTCAACTGGCGGTGAAACAAAGCTCAAACTCACTCCA 9281
Qy 2981 LeuProGluAlaArgLeuLeuAspLeuSerSerTyrPheThrValGlyAlaGlyGlyGly 3000
Db 9282 TTGCGGAGGCGGAGCGCGCTGGATTTATTCGGGTGTTTACCGTGGCGCGCGGGGCG 9341
Qy 3001 AspIleTyrHisSerValSerArgAlaArgProArgLeuLeuLeuPheGlyLeuLeuLeu 3020
Db 9342 GACATTTATCACAGCGTGTGCGATGTCGACCCCGCTTATCTCTTTGCTTACTCTCTA 9401
Qy 3021 LeuPheValGlyValGlyLeuPheLeuLeuProAlaArg 3033
Db 9402 CTTAGCGTAGGAGTAGGCATCTTTTACTCCCGCTCGG 9440
```

RESULT 4

US-07-925-695-6

; Sequence 6, Application US/07925695

; Patent No. 542B145

; GENERAL INFORMATION:

; APPLICANT: OKAMOTO, Hiroaki

```
; APPLICANT: NAKAMURA, Tetsuo
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOME,
; POLYNUCLEOTIDES, POLYPEPTIDES, ANTIGEN, ANTIBODY AND
; DETECTION SYSTEMS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Beveridge, DeGrandi, Weilacher & Young
; STREET: 1850 M Street, N.W., Suite 800
; CITY: Washington
; STATE: D.C.
; COUNTRY: US
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/925,695
; FILING DATE: 19920807
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 287402/91
; FILING DATE: 09-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 360441/91
; FILING DATE: 05-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Weilacher, Robert G.
; REGISTRATION NUMBER: 20,531
; REFERENCE/DOCKET NUMBER: 06/87-48009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2811
; TELEFAX: (202) 659-1462
; TELEX: WUI 64470
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9511 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-07-925-695-6
```

Alignment Scores:

Pred. No.:	0	Length:	9511
Score:	13805.00	Matches:	2522
Percent Similarity:	90.90%	Conservative:	235
Best Local Similarity:	83.15%	Mismatches:	276
Query Match:	85.64%	Indels:	0
DB:	1	Gaps:	0

US-09-980-559-2 (1-3033) x US-07-925-695-6 (1-9511)

Qy	1	MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGln	20
Db	342	AUGAGCACAAUCCUAAACCCUAAAGAAAAACCAAGAAAAACACAAACCGCCGCACAG	401
Qy	21	AspValLysPheProGlyGlyGlnIleValGlyValTyrIleuLeuProArgArg	40
Db	402	GACGUUAAGUUCGCGGCGCGGAGUUCGUGGCGGAGUUCUUGCGCGCGCAGG	461
Qy	41	GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly	60
Db	462	GGCCCCAGGUUGGUGGCGCGGAGAGAAAGAUUUCYAGCGAUCGCCCGCGGUGA	521
Qy	61	ArgArgGlnProIleProLysAspArgGsrThrGlyLysSerTyrGlyLysProGly	80
Db	522	CGACGCCAGCCCAUCCCGAAAGAUCCGCCGCAAGUCCUGGGGAAAGCCAGGA	581
Qy	81	TyrProTyrProLeuTyrGlyAsnGluGlyLeuGlyTyrAlaGlyTyrIleuLeuSerPro	100
Db	582	UAUCCUUGGCCCGUACGGAAGAGGUGCGGCGGUGGCGGUGGCGGUGGCGGUGGCGGCGG	641

QY	101	ArgGlySerArgProSerTyrGlyProAsnAspProArgHisArgSerArgAsnValGly	120	Db	1722	GGGCGGACGAGYUUYCGCAUCGGCGGGGAAACUUUGGAUAACGAACCAACCGUCCAC	1781
Db	642	CGCGGGUCUGUCUACUUGGGGCCCCACCGACCCCGCGCAUGAUCACGCAUUGGGC	701	QY	481	ProGluAspMetArgProTyrCysTyrHisTyrProProArgGlnCysGlyValSer	500
QY	121	LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProValVal	140	Db	1782	GAUGGGAGCAUGAGAGCGGCGUACUUGCGGCAUUAACCCCGAGGCCUUGCGGCAUCGUC	1841
Db	702	AGAGUCAUGCAUACCAUACUGUGUGUUGCGGACCUCAUGGGGUAUCCUGUGU	761	QY	501	AlaLysThrValCysGlyProValTyrCysPheThrProSerProValValGlyThr	520
QY	141	GlyAlaProLeuGlyGlyValAlaArgAlaLeuAlaHisGlyValArgValLeuGluAsp	160	Db	1842	GCUAGGCGUUUGCGGACCGGUGUUAUUGUUAUCCUUGAGCCUUGUGUGGCGACC	1901
Db	762	GGCGCCCGGUGGAGGCGGCGCAGAGCUCUGGCGACCGGUGUAGGGUCCUGGAGGAC	821	QY	521	ThrAspArgLeuGlyAlaProThrTyrThrGlyGluAsnGluThrAspValPheLeu	540
QY	161	GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla	180	Db	1902	ACUGAAGAGCGGCGUACCCACUACCGUGGGGGAACCGAGACCGAUGUUCUCCUG	1961
Db	822	GGGAUAAUUACCAACAGGGAUUUAUCCGGUUGUCUUUUUAUUAUUGUUGU	881	QY	541	LeuAsnSerThrArgProProLeuGlySerTyrPheGlyCysThrTyrMetAsnSerSer	560
QY	181	LeuLeuSerCysIleThrThrProValSerAlaGluValLysAsnIleSerThrGly	200	Db	1962	CSAAUAGCACAAGACCCCGGAGGAGCUUGGUUGCGGCGACUUGGAUAGACGGGACU	2021
Db	882	CUUCUGCAUGCGUCACARUGCCAGUGUGUGCAGCAUACCGGAGGAGGAGGAGGAG	941	QY	561	GlyTyrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer	580
QY	201	TyrMetValThrAsnAspCysThrAsnAspSerIleThrTyrGlnLeuGlnAlaVal	220	Db	2022	GGGUUACUAAGACGUGGUGACCAUUGCCGCAUUAAGAAAGACUACAAACAGCACU	2081
Db	942	UACUACGCCACUAUAGUAGUCCAAACAGCAUACCGGAGGAGGAGGAGGAGGAG	1001	QY	581	ThrAspLeuLeuCysProThrAspCysPheArgLysHisProAspThrThrTyrLeuLys	600
QY	221	LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTyrIle	240	Db	2082	CUCGAUUUAUUGUGCCACAGACUGUUAUAGAAAGACCCAGAGUACCUUAUUAAG	2141
Db	1002	CUCUUAUUGUGAGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUG	1061	QY	601	CysGlySerGlyProTyrLeuThrProArgCysLeuIleAspTyrProTyrArgLeuTyr	620
QY	241	ProValSerProAsnValAlaValGlnArgProGlyAlaLeuThrGlnGlyLeuArgThr	260	Db	2142	UGUGAGAGCGGCGCCUUGUUAACUCCAGGUGCCUGGAGUACUCCUUUAAGUUGUG	2201
Db	1062	CAAGUAAACACCCACGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUG	1121	QY	621	HisTyrProCysThrValAsnTyrThrIlePheLysIleArgMetTyrValGlyGlyVal	640
QY	261	HisIleAspMetValMetSerAlaThrLeuCysSerAlaLeuTyrValGlyAspLeu	280	Db	2202	CAUUAUCCUGGACUUAUUAACUACCAUUAUUAAGCGCGGAGUAGUAGGAGGUG	2261
Db	1122	CAGGUGACAGUAGUAGUAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1181	QY	641	GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp	660
QY	281	CysGlyGlyValMetLeuAlaGlnMetPheIleValSerProGlnHisHisTyrPhe	300	Db	2262	GAGCAUCGAUUCUCCGACAGCAUUCACGCGCGGAGUACGCGGAGUACGUGAAGAU	2321
Db	1182	UGCGGGCGGUGAUGAUUUAUGCGGAGGUGUUAUGGUAUACCAACGACCAACUUC	1241	QY	661	ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGluTyrAlaIleLeuPro	680
QY	301	ValGlnAspCysAsnCysSerIleTyrProGlyThrIleThrGlyHisArgMetAlaTyr	320	Db	2322	AGGGAUAGGGGCGAGCAGAGUCCACUUGCGUACUCCUAGAGGCGGUGUCCCA	2381
Db	1242	ACCAAGAGUGCAACUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU	1301	QY	681	CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisGlnAsnIle	700
QY	321	AspMetMetMetAsnTyrSerProThrAlaThrMetIleLeuAlaTyrAlaMetArgVal	340	Db	2382	UGCUUUCUUCUGACUACCAUACCAUACUACUACUACUACUACUACUACUACUACU	2441
Db	1302	GACAUGAUGCTARUCUGUGUCUCCAAUUAUUAUUAUUAUUAUUAUUAUUAUUAU	1361	QY	701	ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrLysTyrIleValArgTyr	720
QY	341	ProGluValIleAspIleSerGlyAlaHisTyrGlyValMetPheGlyLeuAla	360	Db	2442	GUGGACGUGCAGUACUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU	2501
Db	1362	CCCGARUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUG	1421	QY	721	GluTyrValIleLeuLeuPheLeuLeuAlaAspAlaArgValCysAlaCysLeuTyr	740
QY	361	TyrPheSerMetGlnGlyAlaTyrAlaLysValValIleLeuLeuLeuAlaGly	380	Db	2502	GAGUGGUGAUCUCCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU	2561
Db	1422	UAUUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU	1481	QY	741	MetLeuIleLeuLeuGlyGlnAlaGluAlaLeuLeuGlyLysLeuValIleLeuHisAla	760
QY	381	ValAspAlaArgThrHisThrValGlySerAlaAlaGlnThrThrGlyArgLeuThr	400	Db	2562	AUGCUCAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU	2621
Db	1482	GUGGAUGCAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU	1541	QY	761	AlaSerAlaAlaSerCysAsnGlyPheLeuTyrPheValIlePhePheValAlaIleTyr	780
QY	401	SerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTyr	420	Db	2622	GCUAGYUGUGUAGUCCAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU	2681
Db	1542	RGCCUUUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU	1601	QY	781	TyrIleLysGlyArgValValProLeuAlaThrTyrSerLeuThrGlyLeuTyrSerPhe	800
QY	421	HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer	440	Db	2682	UACUUAAGGCGAGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUG	2741
Db	1602	CACAUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU	1661	QY	801	SerLeuLeuLeuAlaLeuProGlnGlnAlaTyrAlaTyrAspAlaSerValHisGly	820
QY	441	LeuPheTyrThrHisSerPheAsnSerCysProGluArgMetSerAlaCysArg	460	Db	2742	CUCCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU	2801
Db	1662	YUGUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU	1721	QY	821	GlnIleGlyAlaAlaLeuLeuValMetIleThrLeuPheThrThrProGlyTyrLys	840
QY	461	SerIleGluAlaPheArgValGlyTyrGlyAlaLeuGlnTyrGluAspAsnValThrAsn	480				

2802	GAACUGGGCGCGCAUAUUAGUAAUUAUACCAUUCUUAUACCCGAGCAUACAAG	2861	ThrLeuLeuSerArgPheLeuTrpTrpLeuCysTyrLeuLeuLeuThrLeuGlyGluAlaMet	860	
841	ThrLeuLeuSerArgPheLeuTrpTrpLeuCysTyrLeuLeuLeuThrLeuGlyGluAlaMet	860	2862	AUCCUCCUGAGCGGUUCAGUGUGGUGGUCUACUGCUGGUCUUGGCGCGAGGCCAG	2921
2862	AUCCUCCUGAGCGGUUCAGUGUGGUGGUCUACUGCUGGUCUUGGCGCGAGGCCAG	2921	861	ValGlnGluTrpAlaProMetGlnValArgGlyGlyArgAspGlyIleIleTrpAla	880
861	ValGlnGluTrpAlaProMetGlnValArgGlyGlyArgAspGlyIleIleTrpAla	880	2922	AUUCAGCAUUGGUUCCCCUGGAGGUCGAGGGGGCGUGACGGGAUCAUCUGGGUG	2981
2922	AUUCAGCAUUGGUUCCCCUGGAGGUCGAGGGGGCGUGACGGGAUCAUCUGGGUG	2981	881	ValAlaIlePheTyrProGlyValValPheAspIleThrLysTrpLeuLeuAlaValLeu	900
881	ValAlaIlePheTyrProGlyValValPheAspIleThrLysTrpLeuLeuAlaValLeu	900	2982	GCUGUCAUUCACCCAGCCGCUUGUUUGAGGUCACGAUAUGGUUGUUGACAUCCUG	3041
2982	GCUGUCAUUCACCCAGCCGCUUGUUUGAGGUCACGAUAUGGUUGUUGACAUCCUG	3041	901	GlyProAlaTyrLeuLeuLysGlyAlaLeuThrArgValProTyrPheValArgAlaHis	920
901	GlyProAlaTyrLeuLeuLysGlyAlaLeuThrArgValProTyrPheValArgAlaHis	920	3042	GGGCCUGCCUACUCCUUAARAGCGUCUCUGACCGGAUACCGUAUUUGUGAGGCCAC	3101
3042	GGGCCUGCCUACUCCUUAARAGCGUCUCUGACCGGAUACCGUAUUUGUGAGGCCAC	3101	921	AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyGlyArgTyrValGlnMet	940
921	AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyGlyArgTyrValGlnMet	940	3102	GCUUUGUCACGAGUGUGUACCCUGGUGAAACACCUCGCRGGGCGUAGGUACAUCAGAUG	3161
3102	GCUUUGUCACGAGUGUGUACCCUGGUGAAACACCUCGCRGGGCGUAGGUACAUCAGAUG	3161	941	AlaLeuLeuAlaLeuGlyArgTyrPheThrGlyThrTyrIleTyrAspHisLeuThrProMet	960
941	AlaLeuLeuAlaLeuGlyArgTyrPheThrGlyThrTyrIleTyrAspHisLeuThrProMet	960	3162	CUGUUAUACCAUAGGAGGAGUAGGACCGGCACTUACAUCUACGACACCCUCCUCCUUUA	3221
3162	CUGUUAUACCAUAGGAGGAGUAGGACCGGCACTUACAUCUACGACACCCUCCUCCUUUA	3221	961	SerAspTrpAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIleIlePhe	980
961	SerAspTrpAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIleIlePhe	980	3222	UCAACUUGGGCGGCCCGGCUUURCGGACCUUGGCNAUCGCCUGGAGGCCUGUGGUUC	3281
3222	UCAACUUGGGCGGCCCGGCUUURCGGACCUUGGCNAUCGCCUGGAGGCCUGUGGUUC	3281	981	SerProMetGluLysLysValIleValTrpGlyAlaGluThrAlaAlaCysGlyAspIle	1000
981	SerProMetGluLysLysValIleValTrpGlyAlaGluThrAlaAlaCysGlyAspIle	1000	3282	AGCCCAUUGGAGAGAAGUCAUUGUGUGGGGCGUGAGACAGUGGCGUGUGAGACAUC	3341
3282	AGCCCAUUGGAGAGAAGUCAUUGUGUGGGGCGUGAGACAGUGGCGUGUGAGACAUC	3341	1001	LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp	1020
1001	LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp	1020	3342	CUGCAUGGCCUCCCGGUCUCCCGAGGCGUAGUGGARGUUCUGUCGCGGCCUGCCGAC	3401
3342	CUGCAUGGCCUCCCGGUCUCCCGAGGCGUAGUGGARGUUCUGUCGCGGCCUGCCGAC	3401	1021	GlyTyrThrSerLysGlyTrpSerLeuLeuAlaProIleThrAlaTyrAlaGlnGlnThr	1040
1021	GlyTyrThrSerLysGlyTrpSerLeuLeuAlaProIleThrAlaTyrAlaGlnGlnThr	1040	3402	GGCUACCUCCACAGGGGUGGAAKCUCCUAGCUCCCAUUAUGUUCUACUCAGCAACUCU	3461
3402	GGCUACCUCCACAGGGGUGGAAKCUCCUAGCUCCCAUUAUGUUCUACUCAGCAACUCU	3461	1041	ArgGlyLeuLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGluGlnAla	1060
1041	ArgGlyLeuLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGluGlnAla	1060	3462	CGUGUCUCCUGGGUGUACUCCUGUGAGCCUACCGGCGCGACAAAUAGAGCAGCGCU	3521
3462	CGUGUCUCCUGGGUGUACUCCUGUGAGCCUACCGGCGCGACAAAUAGAGCAGCGCU	3521	1061	GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly	1080
1061	GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly	1080	3522	GGCAGAGUCCAGGUVUCUGUCCUCCGUCACACAAACUUUCUUGGGGAGCAUCCAUUUGGGC	3581
3522	GGCAGAGUCCAGGUVUCUGUCCUCCGUCACACAAACUUUCUUGGGGAGCAUCCAUUUGGGC	3581	1081	ValLeuTrpThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro	1100
1081	ValLeuTrpThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro	1100	3582	GUCCUUCGACAGUAUUCACGGGCGUGUAUAUAGACCUUGGCGCGCCCAAGGAGCCA	3641
3582	GUCCUUCGACAGUAUUCACGGGCGUGUAUAUAGACCUUGGCGCGCCCAAGGAGCCA	3641	1101	ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTrpProSerProGly	1120
1101	ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTrpProSerProGly	1120	3642	GUCACUCAGAUUACACACGAGCAGAGGGGACCUUGUGGGAUGGCCUAGUCCCCCGGG	3701
3642	GUCACUCAGAUUACACACGAGCAGAGGGGACCUUGUGGGAUGGCCUAGUCCCCCGGG	3701	1121	ThrLysSerLeuGluProCysThrCysGlyAlaValAlaAspLeuTyrLeuValThrArgAsn	1140
1121	ThrLysSerLeuGluProCysThrCysGlyAlaValAlaAspLeuTyrLeuValThrArgAsn	1140	3702	ACUAAUGAUUGGACCCUUGUACCUUGCGGGGCGUAGACCUUACUCCUGUGUACCCGGAAC	3761
3702	ACUAAUGAUUGGACCCUUGUACCUUGCGGGGCGUAGACCUUACUCCUGUGUACCCGGAAC	3761	1141	AlaAspValIleProAlaArgArgGlyAspLysArgGlyValAlaLeuLeuSerProArg	1160
1141	AlaAspValIleProAlaArgArgGlyAspLysArgGlyValAlaLeuLeuSerProArg	1160	3762	GCUGAUGUCAUUCGGUCCGGAGGAAGAUGACCGAGCGGGUGCAUUAUCCUCCGCCAAGG	3821
3762	GCUGAUGUCAUUCGGUCCGGAGGAAGAUGACCGAGCGGGUGCAUUAUCCUCCGCCAAGG	3821	1161	ProLeuSerThrLeuLysGlySerSerGlyGlyProValLeuLeuCysProArgGlyHisAla	1180
1161	ProLeuSerThrLeuLysGlySerSerGlyGlyProValLeuLeuCysProArgGlyHisAla	1180	3822	CCCCUCUACCCUCAAGGAUCAUCCGGAGGCGCCGUGUCUGUWAGGGGACACGCC	3881
3822	CCCCUCUACCCUCAAGGAUCAUCCGGAGGCGCCGUGUCUGUWAGGGGACACGCC	3881	1181	ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle	1200
1181	ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle	1200	3882	GUGGCGUUGUUCAGACGCGCCUGUGUGCCAGGGGUGAGCCAAUUAUUAUGCAUUAUC	3941
3882	GUGGCGUUGUUCAGACGCGCCUGUGUGCCAGGGGUGAGCCAAUUAUUAUGCAUUAUC				

1201	ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro	1220
3942	CCCUGCAGAACUCAGUURUGCCACACGCGACGCCAGUUCUCUGACAAACAGURCGCG	4001
1221	ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyVal	1240
4002	CCAGCUGUGCCCGAGUCUACAGGUGGGUUAUUGCAGCGCAACACAGGACGCGGAAG	4061
1241	SerThrLysValProValAlaTyrAlaGlnGlyTyrLysValLeuValLeuAsnPro	1260
4062	AGCACCAAGGUCCUGCCGCGUAGCCAGUGGGGUAUAAAGUACUGUACUAAAUCCC	4121
1261	SerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerLysAlaHisGlyIleAsnPro	1280
4122	UCUGUCGGCGCACACUUGGUUUUUGGGCCUACAUGUGCAAAAGCCACGGGAUCAAACCU	4181
1281	AsnIleArgThrGlyValArgThrValThrThrGlyAlaProIleThrTyrSerThrTyr	1300
4182	AAUAUCAAAACUGGAGUGCGACCGUUAACACCGGGGACUCUUAUCUACUCCACUUAU	4241
1301	GlyLysPheLeuAlaAspGlyGlyCysAlaGlyGlyAlaTyrAspIleIleIleCysAsp	1320
4242	GGCAGUUUAUUGCAGUAGGAGGUGUGCAGCCGGGUCUUAUGACAUCACUAAUGCGAC	4301
1321	GluCysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGlnAla	1340
4302	GAAUGCCAUUCAGUGGACGCUACUACCAUCUUGGCAUUGGAACAGUCCUUGACCAAGCU	4361
1341	GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProProGlySerValThr	1360
4362	GAGACCGCAGCGCUCAGCUCAGUGUYUUGGCCACAGCCAGCCUCCCGUACGGUGACA	4421
1361	ThrProHisProAsnIleGluGluValAlaLeuGlyGlnGluGlyGluIleProPheTyr	1380
4422	ACUCCCCACAGUACAUAAGAGAGUGGCCUCCUGUGUACAGAGGGCGAGAUCCUUUUUAU	4481
1381	GlyArgAlaIleProLeuSerTyrIleLysGlyArgHisLeuIlePheCysHisSer	1400
4482	GGCAAAAGCUAUUCCCUAGCUUUCACUAAAGGGGGCGACACACUUAUUGCCAUUCA	4541
1401	LysLysLysCysAspGluLeuAlaAlaAlaLeuArgGlyMetGlyLeuAsnSerValAla	1420
4542	AAGAAGAAGUGCGACAGCUCGCGACGCCUCCCGGGGCAYGGGUGUCAUUGCCGUGCA	4601
1421	TyrTyrArgGlyLeuAspValSerValIleProThrGlnGlyAspValValValAla	1440
4602	UACUUAAGGGGUCUGCAGCUCUCCGUUAUACAACUCAAAGAGACGUGGUGUGUUGCC	4661
1441	ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnVal	1460
4662	ACUAGUCCCUAUGACUGGGUACACCGCGCAUUUGACUCYGUCAUCGACUGUAAUGUU	4721
1461	AlaValThrGlnValValAspPheSerLeuAspProThrPheThrIleThrGlnIle	1480
4722	GCAGUCUCUAGAUUGUUGACUUCAGCCUAGACCAACCCUUCACCAUCCACCAUAAACC	4781
1481	ValProGlnAspAlaValSerArgSerGlnArgArgGlyArgThrGlyArgGlyArgLeu	1500
4782	GUCUCUCAGSACGUCUGUCCCGUAGUACACGUAGAGGGGAGAACUGGAGGGGGCGAUUG	4841
1501	GlyIleTyrArgTyrValSerThrGlyGluArgAlaSerGlyMetPheAspSerValVal	1520
4842	GGCRUUUACAGUAUUUUUCCUGAGGYGARRGGCCGUCUGGGAUUUUGCAGCGUAGUG	4901
1521	LeuCysGluCysTyrAspAlaGlyAlaAlaIleTyrTyrGluLeuThrProSerGluThrThr	1540
4902	CVCUCGAGUGCUAUGAUGCCGGGCGACCCUGGUAGAGCUUACUCCUGCUGAGACUACG	4961
1541	ValArgLeuArgAlaTyrPheAsnThrProGlyLeuProValCysGlnAspHisLeuGlu	1560
4962	GUGAGACUCCGGGCGYUUAUUUACAACGCCCCGUUUUGCCGUUAUGCAGACACCCUGAG	5021

[illegible]

7182	UUCCACCGCGCGUGCCUUCUUGGCGCCGUC	CAGACUA	CAAAUCCUGU	UUUGAGCAGACA	7241
2301	TrpLysArgProAspTyrGlnProAlaThrValAlaGlyCysAlaLeuProProAspG	2320			
7242	UGGAAGAGCGCGGCUAUGAACCCACCCACUGUCUAGGCGUGGCCUCC	CCCCCACACU	7301		
2321	LysThrProThrProProProArgArgArgThrValGlyLeuSerGluAspSerIle	2340			
7302	CAAAAGCGAGUGCCUCCACUCCGAGGCGCGCYAAAUUCUGACCCAGACRAUGUG	7361			
2341	GlyAspAlaLeuGlnGlnLeuAlaIleLysSerPheGlyGlnProProSerGlyAsp	2360			
7362	GAGGGGUCCUCAGGAGGUGGUGGACAAAGUUCUAGCCUCCUCCUCCAGAACAAAC	7421			
2361	SerGlyLeuSerThrGlyAlaGlyAlaAlaAspSerGlySerGlnThrProProAspGlu	2380			
7422	UCCGGUCACUCCACUGGAGCGGAUACCGGAGGAGACAUCUGCCAGCAACCCU	CUGACGAG	7481		
2381	LeuAlaLeuSerGluThrGlySerIleSerSerMetProProLeuGluGlyGluLeuGly	2400			
7482	ACUGCGCGUCUCAAAGACGGGGUCACUGUCUCCUGCCUCCUGAGGGAGAGCGGGA	7541			
2401	AspProAspLeuGluProGluGlnValGluProGlnProProProGlnGlnGlyValAla	2420			
7542	GACCCYGCACCGGAGUUGAACCCAGGGGAUCCGCGUCCUCCUUGAGGGGGAGUGUGAG	7601			
2421	AlaProGlySerAspSerGlySerTrpSerThrCysSerGluGluAspAspSerValVal	2440			
7602	GUCAUUGAUUCGGACUCUAAAGUCUGGUGUCCACAGUCUCUGAUCAGAGAGAUUCUGUUAUC	7661			
2441	CysCysSerMetSerTyrSerTrpThrGlyAlaLeuIleThrProCysSerProGluGlu	2460			
7662	UGCUCGCUCAUGUCAUACUCCUGGACGGGGCCCUCAUAACACCAUGUGGGCCGCAAGAG	7721			
2461	GluLysLeuProIleAsnProLeuSerAsnSerLeuLeuArgTyrHisAsnLysValTyr	2480			
7722	GAGAAUUAUCCGAUACAAACCCUCUGAGUAAUUCGCUCAUGCGGUUCCAUAA	YAAAGGUGUAC	7781		
2481	CysThrThrThrLysSerAlaSerLeuArgAlaLysValThrPheAspArgMetGln	2500			
7782	UCCACAAUCCGAGAGUGCCUUCUGAGGGCAAGAGGUGACUUUUGACAGGGUGCAG	7841			
2501	ValLeuAspSerTyrTyrAspSerValLeuLysAspIleLysLeuAlaLysSerLysVal	2520			
7842	GUGCUGGACGCACACUAUGACUCAGUCUUGCAGSAGCUUAAGCGCGCCUCCUAAGGUU	7901			
2521	ThrAlaArgLeuLeuThrMetGluAlaCysGlnLeuThrProProHisSerAlaArg	2540			
7902	RGUGCGAGGCUCCACAGAGAGAAAGCCUGCGCGCUGACCCCGCCCAUCUCCGCGCA	7961			
2541	SerLysTyrGlyPheGlyAlaLysGluValArgSerLeuSerGlyArgAlaValAsnHis	2560			
7962	UCGGAUACGGAUUGGGCAAAAGAGUGGCGCAGCUUAUCCAGGAGGGCCGUUAACCA	8021			
2561	IleLysSerValTrpLysAspLeuLeuGluAspSerGluThrProIleProThrThrIle	2580			
8022	AUCCGGCGGUGUGGGAGGACCCUCCUGAAGACCAACUAUCCCAUUGACACACUAUC	8081			
2581	MetAlaLysAsnGluValPheCysValAspProThrLysGlyGlyLysLysAlaLysArg	2600			
8082	AUGGCUAAAAUAGAGGUGUUCUGCAUGAUCAUCCAAUAAARGGUGGGAAGAACGCGC	8141			
2601	LeuIleValTyrProAspLeuGlyValArgValCysGluLysMetAlaLeuTyrAspIle	2620			
8142	CUCAUCGUUAUACCCCGACCUUGGGGUGAGGUGUGGAAAGUAGGCGCCUUAUGACAUC	8201			
2621	ThrGlnLysLeuProGlnAlaValMetGlyAlaSerTyrGlyPheGlnTyrSerProAla	2640			
8202	RCACAAAGCUCCCAAGCGGAUUAUGGGGCCAUCCUAGGGUCCUAUACUCCCGCA	8261			
2641	GlnArgValGluPheLeuLeuLysAlaTrpAlaGluLysLysAspProMetGlyPheSer	2660			
8262	GAACGGGUGCAUUCUCCUCAAAGCUGGGGAAGUAGAAAGCAUCCAAUGGGGUUCUG	8321			

2661	TyrAspThrArgCysPheAspSerThrValThrGluArgAspIleArgThrGluGluSer	2680
8322	UAUGACACCCCGUCGUUUUAGACUCAAACCGUCACGAGAGGGACAUAAGAACAAGAAUCC	8381
2681	IleTyrArgAlaCysSerLeuProGluGluAlaHisThrAlaIleHisSerLeuThrGlu	2700
8382	AUAUAUACGGGUGUGUUCUCGCCUAAGAACCCAGAACUGUCAUAACUCGUCUACACGAG	8441
2701	ArgLeuTyrValGlyGlyProMetPheAsnSerLysGlyGlnThrCysGlyTyrArgArg	2720
8442	AGACUUUACGUAGAGGGGCCCAUGACACAACAGCAAAAGGCAUAUCCUGCGGUACAGGCGU	8501
2721	CysArgAlaSerGlyValLeuThrThrSerMetGlyAsnThrIleThrCysTyrValLys	2740
8502	UGCCGCGCAAGCGGKGUUUUUCCACCACGACUGGGGAUAUACCAUGACAGAUUUAACAACA	8561
2741	AlaLeuAlaAlaCysLysAlaAlaGlyIleIleAlaProThrMetLeuValCysGlyAsp	2760
8562	GCCUUGCAGGUGUAGCGUGCRGGGAUCUGGAGCCUUGUUAUGUGUGUGUGUGAGAC	8622
2761	AspLeuValValIleSerGluSerGlnGlyThrGluGluAspGluArgAsnLeuArgAla	2780
8622	GACUUGGUGUACUUCUACAGAGGCCAAGGUAAACGAGGAGGACGACGAAACCUAGAGACU	8681
2781	PheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProArgProGluTyr	2800
8682	UUCACGGAGGCUAUGACGAGUAUUUCGCCCCUCCCGGUGACCUUCCAGACCCGGAUAU	8741
2801	AspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaLeuGlyProGlnGly	2820
8742	GACUUGGAGCUUAUAACAUCUGUCUCCUCAAACAGUAUCGUGAGCGUCGACUCUCGGGU	8801
2821	ArgArgArgTyrTyrLeuThrArgAspProThrThrProIleAlaArgAlaAlaTrpGlu	2840
8802	CGCCGCGGUACUUCUAAACAGAGACCCUACACCAUCCAUACCCGAGCUGCUGGGAA	8861
2841	ThrValArgHisSerProValAsnSerTrpLeuGlyAsnIleIleGlnTyrAlaProThr	2860
8862	ACAGUAAGACACUCCCGUCUAAUUUCGUGGCGGCAACAUCACAGUACGCCCCACA	8921
2861	IleTrpAlaArgMetValLeuMetThrHisPheSerIleLeuMetAlaGlnAspThr	2880
8922	AUCUGGGUCCGGAUGGUCAUAUAGACUACAUUUUCUCAUAUAUUGGCCCAGGACACU	8981
2881	LeuAspGlnAsnLeuAsnPheGluMetTyrGlyAlaValTyrSerValSerProLeuAsp	2900
8982	CUGAAACAAAUAUCUCAAUUUGAGAUGUACGGGGCAGUAUAUCUGGUCUAUUCAG	9041
2901	LeuProAlaIleIleGluArgLeuHisGlyLeuAspAlaPheSerLeuHisThrTyrThr	2920
9042	CUACCGGCCAUAUUGAAGGCUACUAGGGCUUGAGCCUUUUCACUACACACUACUCU	9101
2921	ProHisGluLeuThrArgValAlaSerAlaLeuArgLysLeuGlyAlaProProLeuArg	2940
9102	CCCCAGAAACUCUACAGGGUGGAGCAACUCUCAGAAAACUUGGAGCGCCUCCCCUAGA	9161
2941	AlaTrpLysSerArgAlaArgAlaValArgAlaSerLeuIleSerArgGlyGlyArgAla	2960
9162	CGUGGAGAGUCUGGGCGCGUGCGGAGACUUAUCUACUCCGCCCAAGGAGCGAGGCG	9221
2961	AlaValCysGlyArgTyrLeuPheAsnTrpAlaValLysThrLysLeuLysLeuThrPro	2980
9222	GCCAUUUGGCCGCUACCUUCUACUUGGCGGGUAAACAAGACUCAAACUACUCUCA	9281
2981	LeuProGluAlaArgLeuLeuAspLeuSerSerTrpPheThrValGlyAlaGlyGly	3000
9282	UUGCCGAGGCGAGCCGCUUGAUUUUUCGSGGUGUUCACCGGGGCGCGGGGGGCG	9341
3001	AspIleTyrHisSerValSerArgAlaArgProArgLeuLeuLeuPheGlyLeuLeu	3020
9342	GACAUAUAUACAGCGGUGCGUAGCGACCCGCUUAUACUUAUACUUAUUGCUACUCCA	9401

Copied from 09540843 on 05/19/2004

QY 301 ValGlnAspCysAsnCysSerIleTyrProGlyThrIleThrGlyHisArgMetAlaTyr 320
 Db 901 TTACAGGACTGTAACTGCTCAATTTATCCCGGCATGTGTGGGTCAACGATATGGCTGG 960
 QY 321 AspMetMetMetAsnTrpSerProThrAlaThrMetIleLeuAlaTyrAlaMetArgVal 340
 Db 961 GACATGATGATGAACATGGTCCGCAACAGCCCTAGTGGTGTGGCAGTTACTCCGGATC 1020
 QY 341 ProGluValIleAspIleLeuSerGlyValAlaHisTrpGlyValMetPheGlyLeuAla 360
 Db 1021 CCACAGCCGTCGTGGACATGGTGGCGGGGCCCACTGGGGAGTCTCGCGGCTTGGC 1080
 QY 361 TyrPheSerMetGlnGlyAlaTrpAlaLysValValIleLeuLeuAlaGly 380
 Db 1081 TACTATTCCATGCGGGGAACCTGGCTAAGGTTCTGATTTGCTGCTGCTGCTGCTGCTG 1140
 QY 381 ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr 400
 Db 1141 GTTGACGGGGATACCCACAGTGAACAGGGGGGGCCAGCCCAACCAACAGGCTCGTG 1200
 QY 401 SerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTrp 420
 Db 1201 TCCATGTCGCAAGTGGCGGCTCTCAGAAATCCAGCTTATAACCAATGGGAGTTGG 1260
 QY 421 HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer 440
 Db 1261 CACATCAACAGGACTGCCCTGAACTGCAATGACTCTCTCCAGACTGGGTTCTTCCGCGG 1320
 QY 441 LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg 460
 Db 1321 CTGTTCTACACATAGTTTCAACTGTCGGGTGCCAGAGCGCATGGCCAGTGGCCG 1380
 QY 461 SerIleGluAlaPheArgValGlyTyrGlyValAlaLeuGlnTyrGluAspAsnValThrAsn 480
 Db 1381 ACCATTGACAAAGTTCACACAGGATGGGGTCCCACTACTTATCTGCTAG-----TCTAGC 1434
 QY 481 ProGluAspMetArgProTyrCysTrpHisTyrProProArgGlnCysGlyValValSer 500
 Db 1435 AGATCAGACAGAGGCGCATATGCTGGCACTACCCACCTCCACAAATGTACCATCGTACCT 1494
 QY 501 AlaLysThrValCysGlyProValTyrCysPheThrProSerProValValGlyThr 520
 Db 1495 GCGTCGGAGGTGTGGCGCCAGTGTACTCTTCCACCCCAAGCCCTGCTGCTGGGGGAGC 1554
 QY 521 ThrAspArgLeuGlyAlaProThrTyrThrTyrGlyGluAsnGlnThrAspValPheLeu 540
 Db 1555 ACCGATCGTTTCGGTGTCCCTAGTATAGATGGGGGAGACGAGACTGACGTGCTGCTG 1614
 QY 541 LeuAsnSerThrArgProProLeuGlySerTrpPheGlyCysThrTrpMetAsnSerSer 560
 Db 1615 CTCAACACAGCGCGCGCGCAAGGCAACTGGTTCGGCTGCACATGGATGAATGATGACACC 1674
 QY 561 GlyTyrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer 580
 Db 1675 GGGTTTCAACCAAGACATGTGGGGGGCCCGCTGTAAACATCGGGGGGTTCGGCAAC----- 1728
 QY 581 ThrAspLeuLeuCysProThrAspCysPheArgLysHisProAspThrThrTyrLeuLys 600
 Db 1729 AACACCTGACCTGCCCGCCAGGACTGCTTCCGAGACACCCCGAGGCTACCTACACAAA 1788
 QY 601 CysGlySerGlyProTrpLeuThrProArgCysLeuIleAspTyrProTyrArgLeuTrp 620
 Db 1789 TGTGGTTGGGGCTTGGCTGACACCTAGTGTGATGTTGACTATCCATACAGGCTCTGG 1848
 QY 621 HisTyrProCysThrValAsnTyrThrIlePheLysIleArgMetTyrValGlyGlyVal 640
 Db 1849 CATTTACCCCTGCACTGTAACTTTTACCATCTTCAAGGTTAGGATGTATGTGGGGGGGTG 1908
 QY 641 GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp 660
 Db 1909 GAGCAGGCTCAATGCTGATCAATTTGGACCCGAGGAGAGGATGACTCTCTGGGCCCGCGCAT 1968
 QY 661 ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGluTrpAlaIleLeuPro 680

Db 1969 AGGGATAGGCGGAGGCTCAGCCGCTGCTGCTCTACAAACAGAGTGGGAGGTACTGCGC 2028
 QY 681 CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisGlnLeuHisGlnLeu 700
 Db 2029 TGTTCCTTCCACCACTTACAGCTCTGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2088
 QY 701 ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuLeuThrLysTyrIleValArgTrp 720
 Db 2089 GTGGAGCTGCAATACCTATACGTTATAGGCTCAGCGGTTGCTCTCTTTCCTCAATCAATG 2148
 QY 721 GluTrpValIleLeuLeuPheLeuLeuAlaAspAlaArgValCysAlaCysLeuTrp 740
 Db 2149 GAGTATGCTCTGCTGCTTCTCTCTAGGCGACGCTGCTGCTGCTGCTGCTGCTGCTG 2208
 QY 741 MetLeuIleLeuLeuGlyGlnAlaGluAlaLeuGluLeuValIleLeuHisAla 760
 Db 2209 ATGATGCTCTGATAGCCCGGCGGCGGCTTGGAGAACCTGGTGGTCTCTCAATTCG 2268
 QY 761 AlaSerAlaAlaSerCysAsnGlyPheLeuTyrPheValIlePhePheValAlaAlaTrp 780
 Db 2269 GGTCTGTGGCGGCGCACATGTCCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2328
 QY 781 TyrIleLysGlyArgValProLeuAlaThrTyrSerLeuThrGlyLeuTrpSerPhe 800
 Db 2329 TACATCAAGGAGGCTGCTGCTGCGGCGACATATGCTCTTATGCGCTGCTGCGCTG 2388
 QY 801 SerLeuLeuLeuLeuAlaLeuProGlnAlaTyrAlaTyrAspAlaSerValHisGly 820
 Db 2389 CTCCTGCTCTGCTGCAATTACCGCGAGCTTACGCCATGAGCCGGAGATGGCTGCA 2448
 QY 821 GlnIleGlyAlaAlaLeuValMetIleThrLeuPheThrLeuThrProGlyTyrLys 840
 Db 2449 TCGTGGGAGGCGGCTTTTGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2508
 QY 841 ThrLeuLeuSerArgPheLeuTrpTrpLeuCysTyrLeuLeuThrLeuGlyGluAlaMet 860
 Db 2509 GTGTTCTCTCGCTAGGCTCATATGCTGTATCAATATTTTACCACAGAGCGGAGCGGAC 2568
 QY 861 ValGlnGluTrpAlaProProMetGlnValArgGlyGlyArgAspGlyIleIleTrpAla 880
 Db 2569 TTACATGTGTGATCCCCCTCAACGCTCGGGGAGCGCGGATGCTCATCTCTCTCTC 2628
 QY 881 ValAlaIlePheTyrProGlyValValPheAspIleThrLysTrpLeuLeuAlaValLeu 900
 Db 2629 ATGTGGCAGTCCATCCAGAGCTATCTTTGACATCACCAAACTCTAATGCTGCTACTC 2688
 QY 901 GlyProAlaTyrLeuLeuLysGlyAlaLeuThrArgValProTyrPheValArgAlaHis 920
 Db 2689 GGTCCGCTCATGCTGCTCCAAAGCTGGCATAACAGAGTGGCTGCTGCTGCTGCTGCTG 2748
 QY 921 AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyGlyArgTyrValGlnMet 940
 Db 2749 GGGCTCATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2808
 QY 941 AlaLeuLeuAlaLeuGlyArgTrpThrGlyThrTyrIleTyrAspHisLeuThrProMet 960
 Db 2809 GCTTCATGAAGCTGGCGCGCTGACAGCAGCATATTACCACTTACCACTTACCTGCTA 2868
 QY 961 SerAspTrpAlaAlaSerGlyLeuArgAspLeuAlaValAlaGluProIleIlePhe 980
 Db 2869 CGGGATGGCCACCGCGCGGCTTACGAGACCTTGGGTGGCAGTGGAGCGCTGCTGCTGCT 2928
 QY 981 SerProMetGluLysLysValIleValTrpGlyAlaGluThrAlaAlaCysGlyAspIle 1000
 Db 2929 TCCGATGAGACCAAGATCATCTGCGGAGCAGACACCGCGCGCTGCTGCTGCTGCTGCT 2988
 QY 1001 LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp 1020
 Db 2989 ATCTTGGGTCTCCCGCTCTCCCGCGAGGAGGAGAGATGACTCTCTGGGCCCGCGCAT 3048
 QY 1021 GlyTyrThrSerLysGlyTrpSerLeuLeuAlaProIleThrAlaTyrAlaGlnThr 1040

3049 AGTCTTTAAAGGGCGGGGTTGGACCTCTCGGCCCATCACGGCCTTACTCCCAACAGACG 3108
 1041 ArgGlyLeuLeuGlyThrIleValValSerMetThrGlyArgAspLeuThrGluGlnAla 1060
 3109 CGGGGCTTACTTGGTTGTCATCATCTAGCTTACAGGCGCGGACAAAGAACACAGGTTCGAG 3168
 1061 GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly 1080
 3169 GGAAGGTTTCAGGTGGTTCCACCGCAACACATCTTCTTGGCGACCTCGGTCAACGGC 3228
 1081 ValLeuTrpThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro 1100
 3229 GTGTGTGGACGTTTACCATGTGTCTGGCTCAAGACCTTAGCCGCGCAAAAGGGGCCA 3288
 1101 ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTrpProSerProProGly 1120
 3289 ATCACCAGATGTACACTAATGTGGACCGAGACCTCTCGGTGGCCCAAGCCCCCGGG 3348
 1121 ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArgAsn 1140
 3349 GCGGTTCTTGACCATGACCTGTGGAGCTGACACCTTACTTGTGTACAGACAT 3408
 1141 AlaAspValIleProAlaArgArgGlyAspLysArgGlyAlaLeuLeuSerProArg 1160
 3409 GCTGACGTCTTCGGGTGCGCGCGGGCGCAGTAGGGGAGCGTGTCTCTCCCGCAGG 3468
 1161 ProLeuSerThrLeuLysGlySerSerGlyGlyProValLeuLysProArgGlyHisAla 1180
 3469 CCTGTCTCTTACGTGAAGGGCTTTCGGGTGGTCCACTGTCTGCCCCCTTCGGGGCAGCT 3528
 1181 ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle 1200
 3529 GTGGCATCTTCGGGTGCGGTATGACCCCGGGGGTTCGGAAGCGGTGACTTGTG 3588
 1201 ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspSerThrPro 1220
 3589 CCGGTAGTCCCATGAACTACTATGCGGTCTCCGGTCTTACGGACAACTCATCCCC 3648
 1221 ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLys 1240
 3649 CCGGCCGTACCGCAGTCTTCAAGTGGCCCTACACGCTCCCACTGGCAGCGCAAG 3708
 1241 SerThrLysValProValAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnPro 1260
 3709 AGTACTAAAGTCCGGTGCATATGACGCCAAGGGTACAGGTGTCTGCTCTCATCCG 3768
 1261 SerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerLysAlaHisGlyIleAsnPro 1280
 3769 TCGTTGCCGTACCTTAGGTTTGGGGCGTATATGCTAAGGCACACAGGTATTGACCCC 3828
 1281 AsnIleArgThrGlyValArgThrValThrThrGlyAlaProIleThrTyrSerThrTyr 1300
 3829 AACATCAGAACTGGGGTAAGGACCATTTACACAGCGCCCCCGCTCATCTACTTACTAT 3888
 1301 GlyLysPheLeuAlaAspGlyGlyCysAlaGlyValAlaTyrAspIleIleIleCysAsp 1320
 3889 GGCAGTTTCTTGGCGAGTGTGTCTTGGGGCGCTTATGACATCAATATATGTAT 3948
 1321 GluCysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGlnAla 1340
 3949 GAGTGCCATTCAACTGACTCGACTACAACTTGGGCATCGGCACAGTCTCTGGACCAAGCG 4008
 1341 GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProProGlySerValThr 1360
 4009 GAGACGGCTGGAGCGCGCTTGTCTGCTCGCACCGCTACGCTCCGGGATCGGTACCC 4068
 1361 ThrProHisProAsnIleGluGluValAlaLeuGlyGlnGluGlyGluIleProPheTyr 1380
 4069 GTGCCACACCAACATCGAGAGGTGGCTGTCTAATACTGAGAGATCCCTCTCTAT 4128
 1381 GlyArgAlaIleProLeuSerTyrIleLysGlyGlyArgHisLeuIlePheCysHisSer 1400
 4129 GGCAAGCCATCCCATTTGAAGCCATCAGGGGGGAAGGCATCTCATTTTCTGTCTATCC 4188

1401 LysLysLysCysAspGluLeuAlaAlaLeuArgGlyMetGlyLeuAsnSerValAla 1420
 4189 AAGAAGAAGTGCAGAGCTCGCCCAAGCTGTCCAGGCTCGGAATCAACCTGTGGCG 4248
 1421 TyrTyrArgGlyLeuAspValSerValIleProThrGlnGlyAspValValValAla 1440
 4249 TATTACCGGGGCTCGATGTGTCTCATACCACTATCCGAGAGCTGTCTGTCTGGCA 4308
 1441 ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnVal 1460
 4309 ACAGACGCTCTCATACGGGCTATACGGCGCATTTTGACTCAGTGATCGACTGTAAACA 4368
 1461 AlaValThrGlnValValAspPheSerLeuAspProThrPheThrThrIleThrGlnIle 1480
 4369 TGTGTACCCACAGACAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 4428
 1481 ValProGlnAspAlaValSerArgSerGlnArgArgGlyArgThrGlyArgGlyArgLeu 1500
 4429 GTGCTCAAGACGAGTGTCTCGCTCGCAGCGCGGGGTAGGACTGGCAGGGGTAGGAGA 4488
 1501 GlyIleTyrArgTyrValSerThrGlyGluArgAlaSerGlyMetPheAspSerValVal 1520
 4489 GGCATCTACAGGTTTGTGACTCCGGGAGAACCGCCCTCGGGCATGTTCGATTCTCTCGGTC 4548
 1521 LeuCysGluCysTyrAspAlaGlyAlaAlaTyrTyrGluLeuThrProSerGluThrThr 1540
 4549 CTGTGTGAGTGTATGACCGGGCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4608
 1541 ValArgLeuArgAlaTyrPheAsnThrProGlyLeuProValCysGlnAspHisLeuGlu 1560
 4609 GTTAGTTCGGGCGCTTACCTGAACACACACGAGGTTGCCGTTTGCAGACACCACTGGAG 4668
 1561 PheTrpGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThr 1580
 4669 TTTCTGGAGAGTGTCTTACAGGCTTACCCATATAGATGACACTTCTGTGTGTGTGTGTGT 4728
 1581 LysGlnSerGlyGluAsnPheAlaTyrLeuThrAlaTyrGlnAlaThrValCysAlaArg 1600
 4729 AGCAGGACGAGACACACTTCCCTTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4788
 1601 AlaLysAlaProProProSerTrpAspValMetTrpLysCysLeuThrArgLeuLysPro 1620
 4789 GCTCAGGCGCCACCTCCATCATGGGATCAATGTGGAAGTGTCTCATACGGTGTGAACCT 4848
 1621 ThrLeuValGlyProThrProLeuLeuTyrArgLeuGlySerValThrAsnGluValThr 1640
 4849 ACCTGTCACGGGCCAACACCTTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4908
 1641 LeuThrHisProValThrLysTyrIleAlaThrCysMetGlnAlaAspLeuGluValMet 1660
 4909 CTCACCCACCCATACCAACATACATCATGGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4968
 1661 ThrSerThrTrpValLeuAlaGlyGlyValLeuAlaAlaValAlaAlaTyrCysLeuAla 1680
 4969 ACTAGCACCTGGGT 5028
 1681 ThrGlyCysValCysIleIleGlyArgLeuHisIleAsnGlnArgAlaValValAlaPro 1700
 5029 ACAGGAGT 5088
 1701 AspLysGluValLeuTyrGluAlaPheAspGluMetGluGluCysAlaSerArgAlaAla 1720
 5089 GACAGGAGCTTCTTACAGGTTTCGATGAATGAAGAGTGTGTGTGTGTGTGTGTGTGTGTGT 5148
 1721 LeuIleGluGluGlnArgIleAlaGluMetLeuLysSerLysIleGlnGlyLeuLeu 1740
 5149 TACATTCAGCGGAATGAGCTCGCGGAGCAATTCAGCAGAAAGCGTCTCGGGTACTG 5208
 1741 GlnGlnAlaSerLysGlnAlaGlnAspIleGlnProThrValGlnAlaSerTrpProLys 1760
 5209 CAAACAGCCACCAACAGCGGAGGT 5268

[illegible]

7339 AACTCTTTGTCGCCACCATCAATCATGTTATGCCCACCAACATCTCGCAGCGCGCCCTG 7398
 2490 ArgAlaLysValThrPheAspArgMetGlnValLeuAspSerThrVal 2509
 7399 CGCAGAGAGGTACCTTTGACACTGCAAGTCTCGACAGCACCCTACCGGACGCTG 7458
 2510 LeuLysAspIleLysLeuAlaLysValThrAlaArgLeuLeuMetGluGlu 2529
 7459 CTCAGGAGATGAAGCGGAGGCTCCACAGTTAAGCTTAACCTATCCGTAGAGAA 7518
 2530 AlaCysGlnLeuThrProHisSerAlaArgSerLysThrGlyPheGlyAlaLysGlu 2549
 7519 GCCTGAAGCTGACGCCGCCACATTCGGCCAAATCCAGTTTGGCTATCGGGCAAGGAC 7578
 2550 ValArgSerLeuSerGlyArgAlaValAsnHisIleLysSerValThrLysAspLeuLeu 2569
 7579 GTCCGAACCTATCCAGAGAGCGGCTTAACCATCCATCCGTGTGGAAGGACTTGGCTG 7638
 2570 GluAspSerGluThrProLysThrIleMetAlaLysAsnGluValPheCysVal 2589
 7639 GAAGACACTGTGACACCAATTGACACCATCATGCGCAAAATGAGGTTTCTGTGC 7698
 2590 AspProThrLysGlyLysLysAlaLysValThrAlaArgLeuLeuValThrProAspLeuGlyVal 2609
 7699 CAACACAGAGAGGAGCGGCTTAAGCAGCGCGCTTATCGTATTCAGATCTGGGAGTC 7758
 2610 ArgValCysGluLysMetAlaLeuThrAspIleThrGlnLysLeuProGlnAlaValMet 2629
 7759 CGGTATGCGAGAGAGTGGCCCTCTATGATGTGCTCCACCCCTTCTCCAGTGTGGATG 7818
 2630 GlyAlaSerThrGlyPheGlnThrSerProAlaGlnArgValGluLeuLeuLysAla 2649
 7819 GGCCTCTACAGATTCAGTACTCTCTGGCAGCAGGTCAGTCTCTGGTGAATACC 7878
 2650 TrpAlaGluLysAspProMetGlyPheSerThrAspThrArgCysPheAspSerThr 2669
 7879 TGGAAATCAAGAAAAACCCATCGGCTTTCATATGACACTCGCTGTTTCGACTCAAGC 7938
 2670 ValThrGluArgAspIleThrGluSerIleThrArgAlaCysSerLeuProGlu 2689
 7939 GTCCAGCAGAACACATCGGTGTGAGAGTCAATTTACCAATGTTGACTTGGCCCCC 7998
 2690 GluAlaHisThrAlaIleHisSerLeuThrGluArgLeuThrValGlyGlyProMetPhe 2709
 7999 GAAGCCAGACAGCCATAAAATCGCTCACAGACGCGCTTATATCGGGGTCTCTGACT 8058
 2710 AsnSerLysGlyGlnThrCysGlyThrArgArgCysArgAlaSerGlyValLeuThrThr 2729
 8059 AATTCAAAAGGCGAGAACTCGGCTTATCGCGGTGCGCGCAGCGCGCTGTGACGACT 8118
 2730 SerMetGlyAsnThrIleThrCysThrValLysAlaLeuAlaCysLysAlaAlaGly 2749
 8119 AGTCGGGTAAACCTTACATGTACTTGAAGGCTCTGACGCTGTGAGCTGCGAAG 8178
 2750 IleAlaProThrMetLeuValCysGlyAspLeuValLysSerGluSerGln 2769
 8179 CTCAGGACTGACGATGCTCTGACGAGAGACGCTCTGCTTATCTGTAAGAGCGC 8238
 2770 GlyThrGluAspGluArgAsnLeuAlaPheThrGluAlaMetThrArgThrSer 2789
 8239 GGAACCAAGAGAGCGCGGAGCTTACGAGTCTTCCAGGAGGCTATGACTAGGTACTCC 8298
 2790 AlaProGlyAspProArgProGluThrAspLeuLeuIleThrSerCysSer 2809
 8299 GCCCCCCCGGAGCCGCCAACCAAGATACGACTTGGAGCTGATAACATCATGTTC 8358
 2810 SerAsnValSerValAlaLeuGlyProGlnGlyArgArgThrValLeuThrArgAsp 2829
 8359 TCCATGTGTGCTGCGCCACGATGATCAGGCAAAAGGCTGACTACCTACCCCGTAT 8418
 2830 ProThrThrProIleAlaArgAlaAlaTrpGluThrValArgHisSerProValAsnSer 2849
 8419 CCCACACCCCTAGCAGCGGCTGCTGGGAGACAGCTAGACACACTCCAGTTAACTCC 8478

2850 TrpLeuGlyAsnIleGlnThrAlaProThrIleTrpAlaArgMetValLeuMetThr 2869
 8479 TGGCTAGCAACATTTATGATGATGCGCCACTTGTGGCAAGGATGATTTGATGACT 8538
 2870 HisPheSerIleLeuMetAlaGlnAspThrLeuAspGlnAsnLeuAsnPheGluMet 2889
 8539 CACTTCTTCTCCATCTTCTAGCGAGCAACTTGAAGAGCCCTGACTGCCAGATC 8598
 2890 TyrGlyAlaValThrSerValSerProLeuAspLeuProAlaIleIleGluArgLeuHis 2909
 8599 TAGCGGCGCTGTACTTCCATTCAGCCACTTGACCTCAGATCATTCAGACGACTCCAT 8658
 2910 GlyLeuAspAlaPheSerLeuHisThrThrProHisGluLeuThrArgValAlaSer 2929
 8659 GGCCTTAGCGCATTTTCACTCATGATGATCTCCAGGTGAGATCAATAGGGTGGCTTCA 8718
 2930 AlaLeuArgLysLeuGlyAlaProProLeuArgAlaTrpLysSerArgAlaVal 2949
 8719 TGCCTCAGAAACTTGGGTACACCTTTCGAGTCTGGAGACATCGGCGCAGGCGTC 8778
 2950 ArgAlaSerLeuLysSerArgGlyArgAlaAlaValCysGlyArgThrLeuPheAsn 2969
 8779 CGCGCTAGGCTACTGTCAGGAGGAGGCGCCACTTGTGGCAAAATACCTCTTCAAC 8838
 2970 TrpAlaValLysThrLysLeuLysLeuThrProLeuProGluAlaArgLeuLeuAspLeu 2989
 8839 TGGCGAGTAAACCAACTTAACCTCACTCAATCCCGCTGCGTCCCGGCTGGACTTG 8898
 2990 SerSerTrpPheThrValGlyAlaGlyGlyAspIleThrHisSerValSerArgAla 3009
 8899 TCCGCGTGTGCTGCTGTTACAGCGGGGAGACATATATCACAGCCTGTCTCGTGC 8958
 3010 ArgProArgLeuLeuLeuPheGlyLeuLeuLeuPheValGlyValGlyLeuPheLeu 3029
 8959 CGACCCCGTGTGTTACTGCTGCTGCTACTTCTTCTGTAGGGGTAGGCATCTACCTG 9018
 3030 LeuProAlaArg 3033
 9019 CTCGCCAACCGA 9030

RESULT 6
 US-08-384-616-13
 ; Sequence 13, Application US/08384616
 ; Patent No. 5847101
 ; GENERAL INFORMATION:
 ; APPLICANT: OKAYAMA, Hiroto
 ; APPLICANT: FUKU, Isao
 ; APPLICANT: MORI, Chisato
 ; APPLICANT: TAKAMIZAWA, Akahisa
 ; APPLICANT: YOSHIDA, Iwao
 ; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
 ; NUMBER OF SEQUENCES: 50
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
 ; ADDRESS: Naughton
 ; STREET: 1725 K St. N.W. Suite 1000
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20006
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/384,616
 ; FILING DATE:
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/769,996

FILING DATE: 02-OCT-1991
 APPLICATION NUMBER: JP 2-167466
 FILING DATE: 25-JUN-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 2-230921
 FILING DATE: 31-AUG-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 2-305605
 FILING DATE: 09-NOV-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/635,451
 FILING DATE: 28-DEC-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Stevens-Smith, Theresa M.
 REGISTRATION NUMBER: 36,281
 REFERENCE/DOCKET NUMBER: 900703B
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 659-2930
 TELEFAX: (202) 887-0357
 TELEX: 440142

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:
 LENGTH: 9030 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA from genomic RNA
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 1..9030

OTHER INFORMATION: /note: "sequence = 333 - 9362 of
 OTHER INFORMATION: SEQ ID NO: 1"

FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..9030

US-08-384-616-13

Alignment Scores:

Score: 0 Length: 9030
 Pred. No.: 11984.50 Matches: 2175
 Percent Similarity: 83.02% Conservative: 352
 Best Local Similarity: 71.45% Mismatches: 472
 Query Match: 74.35% Indels: 45
 DB: 2 Gaps: 8

US-09-980-559-2 (1-3033) x US-08-384-616-13 (1-9030)

1 MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgProGln 20
 1 ATGAGCAGCAATCCTAAACCTCAAGAAAAACCAACGTAACACCAACCGCCGCCACAG 60
 21 AspValLysPheProGlyGlyGlnIleValGlyValLysValLysLeuProArgArg 40
 61 GACGTCAAGTTCCTCCGGCGGTGGTTCAGATCGTTGGTGGAGTTTACCTGTTCGGCGCAGG 120
 41 GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly 60
 121 GSCCCAGGTGGTGTGGCGGCCCGCAGAGACTTCGAGCGGTTCGCAACCTCTGTGA 180
 61 ArgArgGlnProLysPheArgArgSerThrGlyLysSerTrpGlyLysProGly 80
 181 AGSCGCAACCTATATCCCAAGGCTCCGCCGCGGCGGAGGACCTGGGTCTAGCCCGCGG 240
 81 TyrProTrpProLeuTyrGlyAsnGluGlyLeuGlyTrpAlaGlyTrpLeuLeuSerPro 100
 241 TACCTCTGGCTCTCTATGGCAATGAGGGCTTAGGGTGGCGAGGATGGCTCTGTCAACC 300
 101 ArgGlySerArgProSerTrpGlyProAsnAspProArgHisArgSerArgHisValGly 120
 301 CGCGGCTCCCGGCTAGTTGGGGGCCCCACGAGACCCCGCGGTAGGTCGGGTAAATTGGGT 360
 121 LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProValVal 140

361 AAGGTGTCATCGATACCTCACAATGCGGCTTCGCCGATCTCATGGGTACATTCGCTCGTC 420
 141 GlyAlaProLeuGlyValAlaArgAlaLeuAlaHisGlyValArgValLeuGluAsp 160
 421 GCGGCCCTCCCTGGGGGGCGCTCCAGGGCCCTGGCACATGGTGTTCGGGTTCCTGGGAC 480
 161 GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla 180
 481 GCGGTGAATATGCAACAGGAATCTGCCGGTGTCTCTTTTCTATCTCTCTCTGGCT 540
 181 LeuLeuSerCysIleThrThrProValSerAlaAlaGluValLysAsnIleSerThrGly 200
 541 CTGCTGTCTGCTCCTCACCACCCAGCTTCCTGCTTACGAAGTGCACAACAGTGCCTGGATA 600
 201 TyrMetValThrAsnAspCysThrAsnAspSerIleThrTrpGlnLeuGlnAlaVal 220
 601 TATCATGTTCAGAACGACTGCTCCAAACGCAACCATTTGTATGAGGACGCGACTTGATC 660
 221 LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTrpIle 240
 661 ATGCATACTCTCTGGGTGGTGGCTTCCTGCTTCGGGAAGCAACTCTCTCCGCTCTGGTA 720
 241 ProValSerProAsnValAlaValGlnArgProGlyAlaLeuThrClnGlyLeuArgThr 260
 721 GCGCTCACTCCCACTGCTGAGCCAGCAAGCTCACCATCCCAACGACGATACGACGC 780
 261 HisIleAspMetValValMetSerAlaThrLeuCysSerAlaLeuTyrValGlyAspLeu 280
 781 CACGTGATCTCTGTTGGGGCGGCTGCTTCTGCTTCGCTATGTACGTGGGGACCTC 840
 281 CysGlyGlyValMetLeuAlaGlnMetPheIleValSerProGlnHisIleSerTrpPhe 300
 841 TGGCGATCTGTTTCTCTCTGCTCTCAGCTGTTTCACTTCTCGCTTCGCCGCGCATGTGACA 900
 301 ValGlnAspCysAsnCysSerIleTyrProGlyThrIleThrGlyHisArgMetAlaTrp 320
 901 TTACAGGACTGTAACTGCTCAATTTATCCCGGCCATGTGTCGGGTACACCATATGGCTGG 960
 321 AspMetMetMetAsnTrpSerProThrAlaThrMetIleLeuAlaTyrAlaMetArgVal 340
 961 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
 341 ProGluValIleIleAspIleSerGlyAlaHisTrpGlyValMetPheGlyLeuAla 360
 1021 CCACAAGCGCTGCTGACATGCTGGCGGGGCCCACTGGGGAGTCTCTGGCGGCGCTTGC 1080
 361 TyrPheSerMetGlnGlyAlaTrpAlaLysValValIleLeuLeuAlaAlaGly 380
 1081 TACTATTCCATGCGCGGAACTGGGCTAAGGTTCTGATTTGTATGCTACTTTTGTCTGTC 1140
 381 ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr 400
 1141 GTTGACGGGATACCACTGACAGGGGGGGCGAAGCCAAACCAACCAACAGGCTCTGTG 1200
 401 SerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTrp 420
 1201 TCCATGTTCCGAAGTGGCGCTCTCAGAAAAATCCAGCTTATAAACACCAATGGAGTGG 1260
 421 HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer 440
 1261 CACATCAACAGGACTGCTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 1320
 441 LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg 460
 1321 CTGTTCTTACACATAGTTTCAACTGCTCGGGTGCAGAGCGCATGCGCCAGTGCCTCC 1380
 461 SerIleGluAlaPheArgValGlyTrpGlyAlaLeuLeuIntyrGluAspAsnValThrAsn 480
 1381 ACCATTGACAGTTTCGACCGGATGGGTCCCATCTACTTATGCTGAG-----TCTAGC 1434
 481 ProGluAspMetArgProTyrCysTrpHisTyrProProArgGlnCysGlyValValSer 500
 1435 AGATCAGACACGAGGCGCATATTGCTGGCACTACTCCACCTCCACCAATGTACCATCGTACCT 1494

QY	501	AlaLeuThrValCysGlyProValTyrCysPheThrProSerProValValValGlyThr	520	QY	861	ValGlnGluTrpAlaProProMetGlnValArgGlyGlyArgAspGlyLeuLeuLeuAla	880
DB	1495	GGGTGGAGTGTGGGGCCAGTGTACTGCTTCAACCCCAAGCCCTGTCTGTGGGAGC	1554	DB	2569	TTACATGTGTGGATCCCCCTCAACGCTCGGGAGGCGCGATGCTCATCTCTCTC	2628
QY	521	ThrAspArgLeuGlyAlaProThrTyrThrTyrGlyGluAsnGluThrAspValPheLeu	540	QY	881	ValAlaLeuPheTyrProGlyValValPheAspIleThrLysTrpLeuLeuAlaValLeu	900
DB	1555	ACCGATCGTTTCGTGTCTACGTATAGATGGGGGAGAACGAGACTGACGTGTGTCTG	1614	DB	2629	ATGTGGCAGTCCATCCAGAGCTAATCTTGTGACATCACAACTTCTAATTCCTCATCTC	2688
QY	541	LeuAsnSerThrArgProProLeuGlySerTyrPheGlyCysThrTrpMetAsnSerSer	560	QY	901	GlyProAlaTyrLeuLeuGlyAlaLeuThrArgValProTyrPheValArgAlaHis	920
DB	1615	CTCAACACACACCGCGCGCAAGCAACTGGTTCGGCTGCACATGGATGAATACACAC	1674	DB	2689	GGTCCGCTCAGTGTCTCAAGCTGGCATACACAGAGTGGCTACTTCTGTGGCGCTCAA	2748
QY	561	GlyTyrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer	580	QY	921	AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyGlyArgTyrValGlnMet	940
DB	1675	GGTTCCACCAAGACATGTGGGGGCCCCGCTGTACATCGGGGGGTGGCAAC-----	1728	DB	2749	GGGCTCATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	2808
QY	581	ThrAspLeuLeuCysProThrAspCysPheArgLysHisProAspThrThrTyrLeuLys	600	QY	941	AlaLeuLeuAlaLeuGlyArgTyrThrGlyThrTyrIleThrLysPheAspHisLeuThrProMet	960
DB	1729	AACACCTGACCTGCCCCACGAGCTGCTTCCGGAAGCAGCCCGAGGCTACCTACACAAA	1788	DB	2809	GCCTTCATCAAGCTGGGCGCTGACAGGACGCTATCATTTACACCATCTTACCCCGCTA	2868
QY	601	CysGlySerGlyProThrLeuThrProArgCysLeuLeuAspTyrProTyrArgLeuTrp	620	QY	961	SerAspTrpAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIleLeuPhe	980
DB	1789	TGTGGTTCGGGCTTGGCTGCACCTAGGTGCTGCTGTCTACACAGAGGAGGCTTGTGAG	1848	DB	2869	CGGATTTGCCACGCGCGGCTTACAGACCTTGGTGGAGTGGAGGCGCTGTCTCTC	2928
QY	621	HisTyrProCysThrValAsnTyrThrIlePheLysLeuArgMetTyrValGlyVal	640	QY	981	SerProMetGluLysLysValIleValTrpGlyAlaGluThrAlaAlaCysGlyAspIle	1000
DB	1849	CATTACCTTGCACCTGTTTACCTTACCTTCAAGTTAGGATGATGTGGGGGGGTG	1908	DB	2929	TCCGACATGGAGACCAAGATCATCATCTGGGGAGCAGACACCGCGCTGTGGGACATC	2988
QY	641	GluHisArgLeuThrAlaAlaCysAsnPheThrArgLysArgCysAsnLeuGluAsp	660	QY	1001	LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp	1020
DB	1909	GAGCACAGCTCAATGCTCATGCAATTGGACCCGAGGAGAGGCTTGTGACTTGGAGGAC	1968	DB	2989	ATCTTGGGCTGTCCCGCTCTCCCGCGAGGAGGAGATCTCTCTGGCGCGCGCGAT	3048
QY	661	ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGluTrpAlaLeuLeuPro	680	QY	1021	GlyTyrThrSerLysGlyTrpSerLeuLeuAlaProIleThrAlaTyrAlaGlnGlnThr	1040
DB	1969	AGGGATAGGCGGAGCTGACCGCTGCTGTCTACACAGAGTGGAGGTACTTGCCC	2028	DB	3049	ACTCTTGAAGGGGGGGTGGAGCTCTCCGCGCCATCACCGCTTACCCACAGAGC	3108
QY	681	CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisGlnAsnIle	700	QY	1041	ArgGlyLeuLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGluGlnAla	1060
DB	2029	TGTTCTTCCACACCTACAGCTCTGTCACTGGCTTGAATCACCTCCATCAGAACATC	2088	DB	3109	CGGGGCTTACTTGTTCATCATCTAGCTTACAGGCGCGGACCAAGAACAGGCTGAG	3168
QY	701	ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrLysTyrIleValArgTrp	720	QY	1061	GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly	1080
DB	2089	GTGAGCTGCAATACCTATAGCTATAGGCTCAGCGTGTCTCTTTCATCAATCAATGG	2148	DB	3169	GGAGAGGTTCAGTGGTTCACACCAACACATCTCTCTCTGGCGACCTGCTCAACGCG	3228
QY	721	GluTrpValIleLeuLeuPheLeuLeuAlaAspAlaArgValCysAlaCysLeuTrp	740	QY	1081	ValLeuTrpThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro	1100
DB	2149	GAGTATGCTGTGTCTTCTCTCTAGCGGACGACGCTGTCTGTCTGTCTGTCTGTCTG	2208	DB	3229	GTGTGTGACCTTTTACCTGCTGCTCAAGAGCTTACAGAGCTTACCGCGCAAGGGCCCA	3288
QY	741	MetLeuIleLeuLeuGlyGlnAlaGluAlaLeuGluLysLeuValIleLeuHisAla	760	QY	1101	ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTrpProSerProGly	1120
DB	2209	ATGATGCTGTATACCCAGCGCGAGGCGCTTGGAGACCTGTGTGTCTCAATTCG	2268	DB	3289	ATCACCAGATGTACACTAATGTGGACAGGACCTCTCTGGCTGGGCCCAAGCCCCCGG	3348
QY	761	AlaSerAlaAlaSerCysAsnGlyPheLeuTyrPheValIlePhePheValAlaAlaTrp	780	QY	1121	ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArgAsn	1140
DB	2269	GGCTGTGTGGCGGCGCATGTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2328	DB	3349	GGCGCTTCTTGCACCATGACCTGTGTGACCTTGTGTGTGTGTGTGTGTGTGTGTGTGT	3408
QY	781	TyrIleLysGlyArgValValProLeuAlaThrTyrSerLeuThrGlyLeuTrpSerPhe	800	QY	1141	AlaAspValIleProAlaArgArgGlyAspLysArgGlyValAlaLeuLeuSerProArg	1160
DB	2329	TACATCAAGGAGGCTGTCTTGTGGGCGACATATGCTCTTATGGCGTGTGGCGCTG	2388	DB	3409	GCTGAGCTCATTCGGGCGCGGGCGGAGACAGTAGGGGAGGCTGTCTCTCCCGAGG	3468
QY	801	SerLeuLeuLeuAlaLeuProGlnAlaTyrAlaTyrAspAlaSerValHisGly	820	QY	1161	ProLeuSerThrLeuLysGlySerSerGlyGlyProValLeuLeuCysProArgGlyHisAla	1180
DB	2389	CTCCTGCTGTGTGGATTTACCCCGGAGCTTACCCATGGACCGGAGATGGGTGCA	2448	DB	3469	CTGTCTCTTCTTGAAGGCTCTTCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3528
QY	821	GlnIleGlyAlaAlaLeuValMetIleThrLeuPheThrLeuThrProGlyTyrLys	840	QY	1181	ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle	1200
DB	2449	TCGTGGGAGGCGGGTTTTTGTGGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2508	DB	3529	GTGGGATCTTCTGGGCTGCGGTATGACCGCGGGGTTCGGAAGCGGTGGACTTTGTG	3588
QY	841	ThrLeuLeuSerArgPheLeuTrpTyrCysTyrLeuLeuLeuGlyGluAlaMet	860	QY	1201	ProValGluThrAspIleValThrArgSerProThrPheSerAspAsnSerThrPro	1220
DB	2509	GTGTTCCTGCTAGGCTCATATGTTGTGTTTACAAATATTTTACCACAGCGGAGCGGAC	2568	DB	3589	CCGTAGAGTTCATGGAACTACTATGCGGTCTTCCGGTCTTTCACGACCACTCATCCCC	3648
QY				QY	1221	ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrThrGlySerGlyLys	1240

[illegible]

4729	AAGCAGCGGAGAGACAACCTTCCCTTACCTGGTAGCATACCAAGGCCACGGTGTGCGCCAGG	4788	1601	AlalysAlaProProSerTripAspValMetTrpLysCysLeuThrArgLeuLysPro	1620
4789	GCTCAGGCCCCACCTCCATCATGGATCAAAATGTGGAAGTGTCTCATACGGCTGAAACCT	4848	1621	ThrLeuValGlyProThrProLeuLeuTrpArgLeuGlySerValThrAsnGluValThr	1640
4849	ACGCTGCACGGGCCAACACCCCTTGCTGTACAGGCTGGGAGCGCTCCAGAAATGAGGTCACC	4908	1641	LeuThrHisProValThrLysTrpIleAlaThrCysMetGlnAlaAspLeuGluValMet	1660
4909	CTCACCCACCCCATCAACCAATACATCATGCATCATGTCCGCTGACCTGGAGGTGCTG	4968	1661	ThrSerThrTrpValLeuAlaGlyValLeuAlaAlaValAlaAlaTrpCysLeuAla	1680
4969	ACTAGCACCTGGGTGCTGGTGGCGGAGTCTTTCGACGCTCTGGCCGCGTATTGCTCTGACA	5028	1681	ThrGlyCysValCysIleIleGlyArgLeuHisIleAsnGlnArgAlaValValAlaPro	1700
5029	ACAGGACGTGGTCATTGTGGTAGGATTATCTTGTCCGGGAGCGCCGGCCATTGTCTCC	5088	1701	AspLysGluValLeuTrpGluAlaPheAspGluMetGluGluCysAlaSerArgAlaAla	1720
5089	GACAGGGAGCTTCTCTACCAAGAGTTGATGAAATGGAAGATGGCGCTCGCACCTCCCT	5148	1721	LeuIleGluGluGlyGlnArgIleAlaGluMetLeuLysSerLysIleGlnGlyLeuLeu	1740
5149	TACATCGAGCAGGAGATCGACGCTGCCGAGCAATTCAGCAGCAAGAGCGCTCGGTTACTG	5208	1741	GlnGlnAlaSerLysGlnAlaGlnAspIleGlnProThrValGlnAlaSerTrpProLys	1760
5209	CAACAGCCACCAACAAAGCGAGGCTGCTGCTCCGTTGGTGGAGTCCAAGTGGCGAGCC	5268	1761	ValGluGlnPheTrpAlaLysHisMetTrpAsnPhelIleSerGlyIleGlnTrpLeuAla	1780
5269	CTTGAGACATTCTGGCGGAAGCACATGTGGAATTTTCATCAGCGGATACAGTACTTAGCA	5328	1781	GlyLeuSerThrLeuProGlyAsnProAlaValAlaSerMetMetAlaPheSerAlaAla	1800
5329	GGCTTATCCATCTGCGCTGGGAAACCCCGCATAGCATTTGNTGGCAATTACACGCTCT	5388	1801	LeuThrSerProLeuSerThrSerThrIleLeuLeuAsnIleLeuGlyGlyTrpLeu	1820
5389	ATCACCGCCGCTCACCCACCAAAAGTACCCTCTCTGTTTAAATCTTGGGGGGTGGGTG	5448	1821	AlaSerGlnIleAlaProProAlaGlyAlaThrGlyPheValValSerGlyLeuValGly	1840
5449	GCTGCCAACCTCGCCCCCCCCAGCGCGCTTTCGGCTTTTCGTGGGCGCGGATCCCGGT	5508	1841	AlaAlaValGlySerIleGlyLeuGlyLysValLeuValAspIleLeuAlaGlyTrpGly	1860
5509	GCGGCTGTGGCAGCATAGGCTCTGGGAGAGTGCTTTGTGGACATTTCTGGCGGTTATGGA	5568	1861	AlaGlyIleSerGlyAlaLeuValAlaPheLysIleMetSerGlyGlyLysProSerMet	1880
5569	GCAGAGTGGCGCGCGCTCGTGCCCTTTAAGGTTCATGAGCGCGAGATGCCCTCCACC	5628	1881	GluAspValValAsnLeuLeuProGlyIleLeuSerProGlyAlaLeuValValGlyVal	1900
5629	GAGGACCTGGTCAATCTACTTCTTGCCATCTCTCTCTCTGGCGCCCTGGTCTCGGGGTG	5688	1901	IleCysAlaAlaIleLeuArgArgHisValGlyProGlyGlyGlyAlaValGlnTrpMet	1920
5689	GTGTGTGCAGCAATACTGCGTCGACACGTGGGTCCGGAGAGGGGGCTGTGCAGTGGATG	5748	1921	AsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaProThrHisTrpValThr	1940
5749	AACCGCTGATAGCTTCCCTCGCGGGGTAAATCATGTTTCCCCCAGCAGCTATGTGCT	5808	1941	GluSerAspAlaSerGlnArgValThrGlnLeuLeuGlySerLeuThrIleThrSerLeu	1960
5809	GAGACGACGCGCGCGCTGTACTCAGATCTCTCCAGCCTTACCATCACTCAGCTG	5868			

QY	1961	LeuArgArgLeuHisAsnTrpIleThrGluAspCysProIleProCysGlyGlySerTrp	1980	QY	2317	ProProProArgLysThrProThrProProArgArgArgThrValGlyLeuSer	2336
DB	5869	CTGAAGAAGGCTCCACAGTGGATTAATGAAGACTGCTCCACACCGTGTCCGGCTCGTG	5928	DB	6949	CCACCTATCAAGGCCCTCCCAATACCACTCCACGAGAAAGAGCGGTTCCTCAACA	7008
QY	1981	LeuArgAspValTrpAspTrpValCysThrIleLeuThrAspPheLysAsnTrpLeuThr	2000	QY	2337	GluAspSerIleGlyAspAlaLeuGlnGlnLeuAlaIleLysSerPheGlyGlnProPro	2356
DB	5929	CTAAGGGATGTTTGGGACTGGATATGACGGTGTGACTGACTTCAAGACCTGGCTCCAG	5988	DB	7009	GAGTCTCGGTCTCTCTGCTTAGCGAGCTCGTACTTAAGACCTTCGGC	7059
QY	2001	SerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTrpLys	2020	QY	2357	ProSerGlyAspSerGlyLeuSerThrGlyAlaGlyAlaAlaAspSerGlySerGlnThr	2376
DB	5989	TCCAGCTCTGCGGAGCTACCTGAGTCCCTTTCTCTGTCGCCAACCGGGTACAAG	6048	DB	7060	-----AGTCCGAATCATCGCGCTCGACAGCGCGACCGGAC	7098
QY	2021	GlyValTrpAlaGlyThrGlyIleMetThrThrArgCysProCysGlyAlaAsnIleSer	2040	QY	2377	Pro-----ProAspGluLeuAla-----LeuSerGluThrGlySerIleSer	2390
DB	6049	GGAGTCTGCGGGGAGCGCATCATGCAAAACACCTGCCCATGTGGACACAGATCAC	6108	DB	7099	GCCTTCTGACGAGCTCCGACGCGGTGACAAAGGATCCGACGTTGAGTCTGACTCC	7158
QY	2041	GlyAsnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTrp	2060	QY	2391	SerMetProProLeuGluGlyGluLeuGlyAspProAspLeuGluProGluGlnValGlu	2410
DB	6109	GGACATGTCAAAACCGTTCATGAGATCGTCGGGCTTAAGACCTGCAGCAACAGTGG	6168	DB	7159	TCCATGCCCTTTCAGGGGAAACCGGGGACCCCGATCTC-----	7200
QY	2061	GlnGlyThrPheProIleAsnCysThrThrGluGlyGlnCysValProLysProAlaPro	2080	QY	2411	ProGlnProProGlnGlyValAlaAlaProGlySerAspSerGlySerTrpSer	2430
DB	6169	CATGGAACATTCCTCCATCAACGATACACCGCGGCCCTGCACACCTCTCCAGCGCA	6228	DB	7201	-----AGTGACGGGCTTGGTCT	7218
QY	2081	AsnPheLysValAlaIleTrpArgValAlaAlaSerGluTrpAlaGluValThrGlnHis	2100	QY	2431	ThrCysSerGluGlu-----AspAspSerValValCysCysSerMetSerTrpSerTrp	2449
DB	6229	AACTATTCTAGGCGCTTGGCGGGTGGCGCTGAGGATGACTGGAGGTACCGGGTG	6288	DB	7219	ACCGTGACGAGGAAAGCTAGTGAGGATGTCTGCTGCTCAATGTCTTACACATGGACA	7278
QY	2101	GlySerTrpHisTrpIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeu	2120	QY	2450	GlyAlaLeuIleThrProCysSerProGluGluGluLysLeuProIleAsnProLeuSer	2469
DB	6289	GGGGATTCCACTACGTGAGGGCATGACCACTGACACAGTAAAGTCCCATGCCAGGTT	6348	DB	7279	GGCGCTTGTGATCAGCCATCGCTCGGAGGAAAGCAAGCTCCCATCAACCGTTGAGC	7338
QY	2121	ProSerProGluPhePheSerTrpValAspGlyValGlnIleHisArgPheAlaProThr	2140	QY	2470	AsnSerLeuLeuArgTrpHisAsnLysValTrpCysThrThrThrLysSerAlaSerLeu	2489
DB	6349	CCGGCTCCTGAATCTCTCGAGGTGGACGAGTGGCTGACAGGTACAGTACGCTCGCG	6408	DB	7339	AACTCTTGTCTGCGCCACCATTAACATGTTTATGCGCAACATCTCGACGCGAGCCCTG	7398
QY	2141	ProLysProPhePheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal	2160	QY	2490	ArgAlaLysLysValThrPheAspArgMetGlnValLeuAspSerTrpTrpAspSerVal	2509
DB	6409	TGACGGCTCTCTACGAGGAGGTTTACATCCAGGTGCGGCTCAACCAATACCTGGTT	6468	DB	7399	CGGACAGAGAAGTCACTTTGACAGACTGCAAGTCTCGACGACCACTACCGGGACGTG	7458
QY	2161	GlySerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThr	2180	QY	2510	LeuLysAspIleLysLeuAlaAlaSerLysValThrAlaArgLeuLeuMetGluGlu	2529
DB	6469	GGGTACAGCTACCATCGAGCCCGAACCAGATGATGATGCTCTCCATCTCCATGCTCACC	6528	DB	7459	CTCAAGAGATGAAGCGAAGCGCTCCACAGTTAAGGCTTAAACTCTATCCGTAGAGGAA	7518
QY	2181	AspProSerHisIleThrAlaGluThrAlaAlaArgArgLeuAlaArgGlySerProPro	2200	QY	2530	AlaCysGlnLeuThrProProHisSerAlaArgSerLysTrpGlyPheGlyAlaLysGlu	2549
DB	6529	GACCCCTCCACATCATCAGCAGAAAGCGCTAAGCTAGGTGGCCAGGGGCTCTCCCCC	6588	DB	7519	GCTCGAAGTACGCGCCCACTTCGGCCAAATCCAGTTTGGCTTATCGGGCAAGGAC	7578
QY	2201	SerGluAlaSerSerAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThr	2220	QY	2550	ValArgSerLeuSerGlyArgAlaValAsnHisIleLysSerValTrpLysAspLeuLeu	2569
DB	6589	TCCTTGCCAGCTCTTCAGCTAGCTAGTGTCTGCGCTTCTTGAAGGCGATGCATC	6648	DB	7579	GTCCGGAACCTATCCAGCAAGCGCTTAACACATCCACTCCGTGTGGAAGGACTTGTCTG	7638
QY	2221	ThrHisGlyLysAlaTrpAspValAspMetValAspAlaAsnLeuPhe-----	2236	QY	2570	GluAspSerGluThrProIleProThrIleMetAlaLysAsnGluValPheCysVal	2589
DB	6649	ACCCACCATCTCTCCGAGCGCTGACCTCATCGAGGCCCACTCTCTGTGGCGGAGGAG	6708	DB	7639	GAAGACACTGTGACCAATTCACACCATCATGGCAAAATGAGGTTTCTGTCTGTC	7698
QY	2237	MetGlyCysValThrArgIleGluSerGlySerLysValValValLeuAspSerLeu	2256	QY	2590	AspProThrLysGlyGlyLysLysAlaAlaArgLeuIleValTrpProAspLeuGlyVal	2609
DB	6709	ATGGGGGGAACATCACCCCGTGGAGTCCGAGAACAAAGGTGTAGTCTCTGGACTCTTC	6768	DB	7699	CAACAGAGAAGGAGCGCGTAAGCCCGCTTATCGTATTCACATCTGGGAGTC	7758
QY	2257	AspProMetValGluGluArgSerAspLeuGluProSerIleProSerGluTrpMetLeu	2276	QY	2610	ArgValCysGluLysMetAlaLeuTrpAspIleThrGlnLysLeuProGlnAlaValMet	2629
DB	6769	GACCCGCTTCGAGGAGGAGGATGAGAGGAAGTATCCGTTCCGGCGGAGATCTTCGGG	6828	DB	7759	CGTGTATGCGAGAAGATGCCCTCTATGATGTGCTCCACCCCTCTCCAGGTCGTGATG	7818
QY	2277	ProLysLysArgPheProAlaLeuProAlaTrpAlaArgProAspTrpAsnProPro	2296	QY	2630	GlyAlaSerTrpGlyPheGlnTrpSerProAlaGlnArgValGluPheLeuLeuLysAla	2649
DB	6829	AAATCCAGAAGTTCCTCCGAGCGATGCCCATCTGGCGCGCCCGGATTACAACCTCCA	6888	DB	7819	GGCTCTCATACGATTCAGTACTCTCTCTGGGACGAGTCCAGTCTCTGGTGAATACC	7878
QY	2297	LeuValGluSerTrpLysArgProAspTrpGlnProAlaThrValAlaGlyCysAlaLeu	2316	QY	2650	TrpAlaGluLysLysAspProMetGlyPheSerTrpAspThrArgCysPheAspSerThr	2669
DB	6889	CTGTAGAGTCTCGGAAGGACCGGACTACGTCTCCCTCGGTGTGACGCGGTGCGCGTTG	6948	DB	7879	TGGAATCAAGAAACCCCATGGCTTTTATATGACACTCGCTGTTTTCGACTCAACG	7938
QY				QY	2670	ValThrGluArgAspIleArgThrGluGluSerIleTyArgAlaCysSerLeuProGlu	2689

```
Db 7939 GTCCAGGAGAACGACATCCGCTGTGAGGAGTCAATTTACCAATGTTGTGACTTTGGCCCC 7998
Qy 2690 GluAlaHisThrAlaIleHisSerLeuThrGluArgLeuTyrValGlyGlyProMetPhe 2709
Db 7999 GAGCCAGACAGCCCAATAAATCGCTCACAGACGGCTTTATATCGGGGTCTCTGACT 8058
Qy 2710 AsnSerIysGlnThrCysGlyTyrArgArgCysArgAlaSerGlyValLeuThrThr 2729
Db 8059 AATTCAAAAGGGCAGAAATCGCGTATTCGCGGTTCGCGCGAGCGGTGTGACGACT 8118
Qy 2730 SerMetGlyAsnThrIleThrCysTyrValIysAlaLeuAlaCysIysAlaAlaGly 2749
Db 8119 AGCTGCGGTAAACCTTCACATCTTACTTGAAGGCTCTGCAGCCTGTGAGCTGGCAG 8178
Qy 2750 IleIleAlaProThrMetLeuValCysGlyAspLeuValIleSerGluSerGln 2769
Db 8179 CTCAGGACTGCGAGGATCTCGTGAACGGAGAGCCTCTGCTTATCTGTGAAGCGCG 8238
Qy 2770 GlyThrGluGluAspGluArgAsnLeuArgAlaPheThrGluAlaMetThrArgTyrSer 2789
Db 8239 GGAACCAAGAGAGCGCGCGAGCTTACGAGTCTTCACGGAGGCTATGACTAGTACTCC 8298
Qy 2790 AlaProProGlyAspProArgProGluTyrAspLeuLeuLeuIleThrSerCysSer 2809
Db 8299 GCCCCCCCGGGGACCGCCCAACCAAGATACGACTTGCAGCTGATACATCATGTTC 8358
Qy 2810 SerAsnValSerValAlaLeuGlyProGlnGlyArgArgTyrTyrLeuThrArgAsp 2829
Db 8359 TCCAAATGTGCGTGCCTCCACAGTATGATCAGGCAAGGGTGACTACCTCACCCTGAT 8418
Qy 2830 ProThrThrProIleAlaArgAlaAlaTrpIleThrValArgHisSerProValAsnSer 2849
Db 8419 CCCACACCCCCCTAGACGGGTGCGTGGAGACAGCTAGACACACTCCAGTTACTCC 8478
Qy 2850 TrpLeuGlyAsnIleIleGlnTyrAlaProThrIleTrpAlaArgMetValLeuMetThr 2869
Db 8479 TGGCTAGGCAACATATTATGATGCGCCCACTTGTGGCAAGGATGATCTGTGACT 8538
Qy 2870 HisPhePheSerIleLeuMetAlaGlnAspThrLeuAspGlnAsnLeuAsnPheGluMet 2889
Db 8539 CACTTCTTCCATCTTCTAGCGAGGAGCAACTTGAAGAACCCCTGGACTCCAGATC 8598
Qy 2890 TyrGlyAlaValTyrSerValSerProLeuAspLeuProAlaIleIleGluArgLeuHis 2909
Db 8599 TAGCGGCGCTGTACTCCATGAGCCACTTGACTTACTTACCTCAGATCATTTGACGACTCCAT 8658
Qy 2910 GlyLeuAspAlaPheSerLeuHisThrTyrThrProHisGluLeuThrArgValAlaSer 2929
Db 8659 GGCCTTAGCGCATTTTCACTCCATAGTACTCTCCAGGTGAGATCAATAGGGTGGCTTCA 8718
Qy 2930 AlaLeuArgIysLeuGlyAlaProProLeuArgAlaTrpIysSerArgAlaAlaVal 2949
Db 8719 TGCTTCAGGAAATCTTGGGTACCACTTTCGAGTCTGGAGACATCGGGCCAGGACGCTC 8778
Qy 2950 ArgAlaSerLeuIleSerArgGlyGlyArgAlaAlaValCysGlyArgTyrLeuPheAsn 2969
Db 8779 CGCGCTAGGCTACTGTCAGGAGGAGGAGCGCCCACTTGTGGCAATACCTTCTTCAAC 8838
Qy 2970 TrpAlaValIysThrIysLeuLeuThrProLeuProGluAlaArgLeuLeuAspLeu 2989
Db 8839 TGGGCAAGTAAACCAAACTTAACTCACTTCAATCCCGCTCGCTGGCTGGACTTG 8898
Qy 2990 SerSerTrpPheThrValGlyAlaGlyGlyAspIleTyrHisSerValSerArgAla 3009
Db 8899 TCCGGCTGGTTCGTTCTGCTGTTACAGCGGGGAGACATATATCAGACCTGTCTCGTCC 8958
Qy 3010 ArgProArgLeuLeuPheGlyLeuLeuPheValGlyValGlyLeuPheLeu 3029
Db 8959 CGACCCCGTGGTTGCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 9018
Qy 3030 LeuProAlaArg 3033
```

```
Db 9019 CTCCCAACCGA 9030
RESULT 7
US-08-904-686A-13
; Sequence 13, Application US/08904686A
; Patent No. 5998130
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chiato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westernman, Hattori, McLeland &
; ADDRESSEE: Naughton
; STREET: 1725 K St., N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/904,686A
; FILING DATE: 01-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: McLeland, Le-Nhung
; REGISTRATION NUMBER: 31,541
; REFERENCE/DOCKET NUMBER: 900703G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9030 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA from genomic RNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..9030
; OTHER INFORMATION: /note: "sequence = 333 - 9362 of
; OTHER INFORMATION: SEQ ID NO: 1"
; NAME/KEY: CDS
```

Copied from 09540843 on 05/19/2004

Db	1969	AGGGATAGCGCGAGCTCAGCCCGCTGCTGCTACACAGAGTGGCAGTACGCTCC	2028	Db	3049	AGTCTTTAAAGGCGGGGTTGCGACTCTCTCGCCCTCATCGCCCTACTCCCAACAGACG	3108
Qy	681	CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisGlnAsnIle	700	Qy	1041	ArgGlyLeuLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGluGlnAla	1060
Db	2029	TGTTCTTCACACCCCTACACAGCTCTGTCTCACTGGCTTGATTCACCTCCATCAGAACATC	2088	Db	3109	CGGGGCTACTTGGTTGCATCATCTAGCTTTACAGCCCGGCGGACAAAGAACAGTGCAG	3168
Qy	701	ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrLysTyrIleValArgTyr	720	Qy	1061	GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly	1080
Db	2089	GTGCACGTGCAATACCTATACGTATAGGGTCAGCGGTGTCTCTTTGCAATCAATGG	2148	Db	3169	GGAGAGGTTGAGTGGTTTCCACCGCAACAACATCTCTCTCGCGACCTCGCTCAACGGC	3228
Qy	721	GluTrpValIleLeuPheLeuLeuLeuAlaAspAlaArgValCysAlaCysLeuTyr	740	Qy	1081	ValLeuTrpThrValTyrHisGlyAlaGlyAsnLysThrIleuAlaGlySerArgGlyPro	1100
Db	2149	GAGTATGCTGCTGTTCTCTCTCTAGCGACACAGTGTCTGTGCTGCTGCTG	2208	Db	3229	GTGTGTGGACCTTTACCATGTGTGCTCAAGACCTTACCGCCGCAAGAGGGGCCA	3288
Qy	741	MetLeuIleLeuLeuGlyGlnAlaGluAlaLeuGluLysLeuValIleLeuHisAla	760	Qy	1101	ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTyrProSerProProGly	1120
Db	2209	ATGATGCTGTATAGCCAGCGCCGCGCCCTTGAGAACCTGTGTGGTCTCTCATTCG	2268	Db	3289	ATCACCCAGATGTACTAATGTGGACAGGACTGTCTGGCTGGCCCAAGCCCCCGGG	3348
Qy	761	AlaSerAlaAlaSerCysAsnGlyPheLeuTyrPheValIlePhePheValAlaAlaTyr	780	Qy	1121	ThrLysSerLeuGluProCysThrCysGlyAlaAlaAspLeuTyrLeuValThrArgAsn	1140
Db	2269	CGGTCTGTGCGCGCGCACATGGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2328	Db	3349	CGCGGTCTCTTGACACCATGCTGCTGGCAGCTCAGACCTTTTACTTGGTTCAGAGACAT	3408
Qy	781	TyrIleLysGlyArgValValProLeuAlaThrTyrSerLeuThrGlyLeuTyrSerPhe	800	Qy	1141	AlaAspValIleProAlaArgArgArgGlyAspLysArgGlyAlaLeuLeuSerProArg	1160
Db	2329	TACATCAAGAGCAGGCTGTCTCTGCGGCGACATATGCTCTTTATGCGGTGTGCGCGCTG	2388	Db	3409	GCTGACGTCTATTCGGTGGCGCGCGCGCGCAGTAGGGGGAGCTGTCTCTCCCCCAGG	3468
Qy	801	SerLeuLeuLeuAlaLeuProGlnGlnAlaTyrAlaTyrAspAlaSerValHisGly	820	Qy	1161	ProLeuSerThrLeuLysGlySerSerGlyProValLeuCysProArgGlyHisAla	1180
Db	2389	CTCTGCTCTGTGCGCATACACCGCGAGCTTACGCCATGAGCGGAGATGGTGCA	2448	Db	3469	CTGTCTCTCTTGAAGGGCTCTTCGGGTGGTCCACTGCTCTGCCCCCTTCGGGCGACGT	3528
Qy	821	GlnIleGlyAlaAlaLeuValMetIleThrLeuPheThrLeuThrProGlyTyrLys	840	Qy	1181	ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle	1200
Db	2449	TCGTGCGGAGCGGGTTTGTGGTCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2508	Db	3529	GTGCGCATCTCTCGGCTGCGGTATGATCACCCCGGGGGTTCGGAAGCGGTGGACTTTGTG	3588
Qy	841	ThrLeuLeuSerArgPheLeuTyrTrpLeuCysTyrLeuLeuThrLeuGlyGluAlaMet	860	Qy	1201	ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro	1220
Db	2509	GTGTCTCTGCTAGGCTCATATGCTGTGTACAAATTTTACCACAGAGCGGAGCGGAC	2568	Db	3589	CCCGTAGAGTCCATGGAACTACTATGCGGTCTCCGGTCTTTCACGGACAACTCATCTCCC	3648
Qy	861	ValGlnGluTrpAlaProMetGlnValArgGlyGlyArgAspGlyIleIleTrpAla	880	Qy	1221	ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLys	1240
Db	2569	TTACATGTGGATCCCCCCCCCTCAACGCTCGGGGAGCGCGCATGCATCTCTCTCTCT	2628	Db	3649	CCGGCGTACCGCAGTCATTTCAAGTGGCCCACTACGCTCCCTCCCTGCGCGGCGGACG	3708
Qy	881	ValAlaIlePheTyrProGlyValValPheAspIleThrLysTrpLeuLeuAlaValLeu	900	Qy	1241	SerThrLysValProValAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnPro	1260
Db	2629	ATGTGCGCAGTCCATCCAGAGCTAATCTTTGACATCACCAAACTTCTAATGGCATATC	2688	Db	3709	AGTACTAAAGTCCGCTGCATATGCAGCCCAAGGTACAAAGTGTCTCTCTCAATCCG	3768
Qy	901	GlyProAlaTyrLeuLeuLysGlyAlaLeuThrArgValProTyrPheValArgAlaHis	920	Qy	1261	SerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerLysAlaHisGlyIleAsnPro	1280
Db	2689	GGTCCGCTCATGTGTCTCAAGCTGGCATTAACAGAGTGGCTTCTGTGCGCGCTCAA	2748	Db	3769	TCCGTTGCGCTACCTTAGGGTTTGGGGGTATATGTCTAAGGCACACGCTATTGACCCC	3828
Qy	921	AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyArgTyrValGlnMet	940	Qy	1281	AsnIleArgThrGlyValArgThrValThrThrGlyValAlaProIleThrTyrSerThrTyr	1300
Db	2749	GGGCTCATGTGATGCTAGTGGGAGAGTCTGCTGGGTCTAATATGTCAAATG	2808	Db	3829	AACTCAGAACTGGGGTNAAGACCATTCACAGGCGCCCGCTCACACTCTTACCTAT	3888
Qy	941	AlaLeuLeuAlaLeuGlyArgTyrThrGlyThrTyrIleTyrAspHisLeuThrProMet	960	Qy	1301	GlyLysPheLeuAlaAspGlyGlyCysAlaGlyGlyAlaTyrAspIleIleCysAsp	1320
Db	2809	GCCTTCATGAAGCTGGGCGGCTGACAGGACAGTACATTTACAACCATCTTACCCGCTA	2868	Db	3889	GGCAAGTTTCTTGGCCGATGTTGCTCTGCGGGCGCTTATGACATCAATATATGAT	3948
Qy	961	SerAspTrpAlaAlaSerGlyLeuArgAspLeuAlaValAlaGluProIleIlePhe	980	Qy	1321	GluCysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGlnAla	1340
Db	2869	CGGATTTGGCCACGCGCGGCTACGAGACCTTGGGTGGCAGTGGAGCGGCTGCTCTC	2928	Db	3949	GAGTGGCATTAACCTGACTGCTACATCTTGGGATCGGCACAGCTCTGGACCAAGCG	4008
Qy	981	SerProMetGluLysLysValIleValTrpGlyAlaGluThrAlaAlaCysGlyAspIle	1000	Qy	1341	GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProProGlySerValThr	1360
Db	2929	TCCGACATGAGACCAAGATCATCCTGGGGAGCAGACACCGCGGCTGTGGGACATC	2988	Db	4009	GAGACGGCTGGAGCGGCTGTGTGTGCTCGCCACCGCTACGCTCCGGATCGGTACC	4068
Qy	1001	LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp	1020	Qy	1361	ThrProHisProAsnIleGluGluValAlaLeuGlyGlnGluGlyGluIleProPheTyr	1380
Db	2989	ATCTTGGGTCTGCCCTCTCGCCCGGAGGGGAAGGAGATCTCTCTGGGCCCGCGCAT	3048	Db	4069	GTGCCACACCAACATCGAGGAGTGGCCCTGTCTAATACTAGAGAGATCCCTCTCTAT	4128
Qy	1021	GlyTyrThrSerLysGlyTyrSerLeuLeuAlaProIleThrAlaTyrAlaGlnThr	1040	Qy	1381	GlyArgAlaIleProLeuSerTyrIleLysGlyGlyArgHisLeuIlePheCysHisSer	1400
Db				Db	4129	GGCAAGCCATCCCATCGAAGCCATCAGGGGGGGAAGGCATCTCTCTCTCTCTCTCTCT	4188

1761	ValGluGlnPheTrpAlaLysHisMetTrpAsnPhelIleSerGlyIleGlnTyrLeuAla	1780
5269	CTTGAGACATTTCTGGCGGAAGCACAATGTGGAAATTTTCATCAGCGGATACAGTACTTAGCA	5328
1781	GlyLeuSerThrLeuProGlyAsnProAlaValAlaSerMetMetAlaPheSerAlaAla	1800
5329	GGCTTTATCCATCTTCCTGGGAAACCCCGCAATAGCATCATTTGATGGCATTCACAGCCTCT	5388
1801	LeuThrSerProLeuSerThrSerThrThrIleLeuLeuAanIleLeuGlyGlyTrpLeu	1820
5389	ATCACCGACCGCTCACCAACAAAGTACCCTCTCTGTTTAAATCATCTTGGGGGGTGGGTG	5448
1821	AlaSerGlnIleAlaProProAlaGlyAlaThrGlyPheValValSerGlyLeuValGly	1840
5449	CTCTGCCAACTCGCCCCCCCCCGCGCTTCGGCTTTTCGTGGCGCGCATATCCCGCT	5508
1841	AlaAlaValGlySerIleGlyLeuGlyLysValLeuValAspIleLeuAlaGlyTyrGly	1860
5509	CGCGCTGTGTGCAGATAGGCTTGTGGAAAGTGCTTGTGGACATTTCTCGCGGTATTATGA	5568
1861	AlaGlyIleSerGlyAlaLeuValAlaPheLysIleMetSerGlyGluLysProSerMet	1880
5569	CGAGAGTGGCGCGCGCTCGTGGCTTTTAAGGTCATGAGCGGCGAGATGCCCTCCACC	5628
1881	GluAspValValAsnLeuLeuProGlyIleLeuSerProGlyAlaLeuValValGlyVal	1900
5629	GAGGACCTGGTCAATCTACTTCTGCCATCCTCTCTCTGGCGCCCTCGTCTCGGGTCT	5688
1901	IleCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrpMet	1920
5689	GTGTGTGAGCAATACTCGTCGTGCACACGTGGTCCGGAGAGGGGGCTGTGCAGTGGATG	5748
1921	AsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaProThrHisTyrValThr	1940
5749	AACCGGTGATAGGTTTCCTCGCGGGGTAAATCATGTTTCCCCACGCACATATGTGCCT	5808
1941	GluSerAspAlaSerGlnArgValThrGlnLeuLeuGlySerLeuThrIleThrSerLeu	1960
5809	GAGAGCGACGCGCGCGCTGTACTCAGATCTCTCTCCAGCCTTACCATCATCATCAGCTG	5868
1961	LeuArgArgLeuHisAsnTrpIleThrGluAspCysProIleProCysGlyGlySerTrp	1980
5869	CTGAAAAGGCTCCACAGTGGATTAATGAAGACTGCTCCACACCCTGTTCCGGCTCGTGG	5928
1981	LeuArgAspValTrpAspTrpValCysThrIleLeuThrAspPheLysAsnTrpLeuThr	2000
5929	CTAAGGATGTTTGGACTGGATATGACGGTGTGACTGACTTCAAGACCTGGCTCCAG	5988
2001	SerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTyrLys	2020
5989	TCCAAGCTCTCGCGCAGTACTCTGGAGTCCCTTTTCTCTGTGCAACGCGGGTACAAAG	6048
2021	GlyValTrpAlaGlyThrGlyIleMetThrThrArgCysProCysGlyValaAsnIleSer	2040
6049	GGAGTCTCGCGGGAGACGGCATCATGCAAAACCACTCGCCCATGTGGAGCACAGATCAC	6108
2041	GlyAsnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTrp	2060
6109	GGACATGTCAAAACGGTTCATGAGGATCTGTGGGGCTTAAGACCTTCAGACACACGTGG	6168
2061	GlnGlyThrPheProIleAsnCysTyrThrGluGlyGlnCysValProLysProAlaPro	2080
6169	CATGAAACATTCCTCCATCAACGCATACACACCGGGCCCCCTGCACACCTCTCTCAGCGCCA	6228
2081	AsnPheLysValAlaIleTrpArgValAlaAlaSerGluTyrAlaGluValThrGlnHis	2100
6229	AATATTCTAGCGCGCTGTGGCGGTGGCCGTGAGGAGTACGTGGAGGTTCACCGGGGTG	6288
2101	GlySerTyrHisTyrIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeu	2120
6289	GGGGATTTCCACTACTGACGGGCATGACCATGTACACAGTAAAGTGCCCATGCCAGTTT	6348
2121	ProSerProGluPhePheSerTrpValAspGlyValGlnIleHisArgPheAlaProThr	2140

D	b	7399	AAC	TCTTTTGTGCGCCACCATAAATCATGGTTATTAGCCAAACAATCTCGCAGCCGACGCCTG	7399
Q	y	2490	ArgAla	lylsyValThrPheAspArgMetGlnValLeuAspSerTyrTyrAspSerVal	2509
D	b	7399	CGG	CAGAAGAAGGTCACTTTTGACAGACTGCAAGTCTTCGACGACCACCTACCGGACGCTG	7458
Q	y	2510	LeuLys	AspIlellysLeuAlaLaSerLysValThrAlaArgLeuLeuThrMetGluGlu	2529
D	b	7459	CTC	AAGGAGATGAAGCGGAAGGCGTCCACAGTTTAAGCGCTAAACTCTCTATCCGTAGAGGAA	7518
Q	y	2530	AlaCy	GlnLeuThrProProHisSerAlaAaSerLysTyGlyPheGlyAlalyeGlu	2549
D	b	7519	GCCT	GCAAGCTGACGCCCCCACATTCGCGCAAATCCAAGTTTTGGCTTAGGGGCAAGAAC	7578
Q	y	2550	ValArg	SerLeuSerGlyArgAlaValAasnHisIleLysSerValTrpLysAspLeuLeu	2569
D	b	7579	GTCC	GGAAACCTATCCAGCAAGGCGGTNACCACATCCACTCCGTGTGGAAAGGACTTGCTG	7638
Q	y	2570	GluAsp	SerGluThrProfileProThrThrIleMetAlaLysAasnGluValPheCysVal	2589
D	b	7639	GAAG	ACACTGTGACCACTTGACACCACTCATGCGCAAAAAAATGAGGTCTTCTGTGTC	7698
Q	y	2590	AspPro	ThrLysGlyLysLysAlaAaArgLeuIleValTyrProAspLeuGlyVal	2609
D	b	7699	CAAC	CAGAAAAGGAGCGGTAAAGCCAGCCGCGCTTATCGTATPTCCAGATCTGGGAGTCT	7758
Q	y	2610	ArgVal	CysGluLysMetAlaLeuTyrAspIleThrGlnLysLeuProGlnAlaValMet	2629
D	b	7759	CGTG	TATGCCAGAAGATGGCCCTCTATGATGTGGTCTCCACCTTCTCAGGTCTGTGATG	7818
Q	y	2630	GlyAla	SerTyrGlyPheGlnTyrSerProAlaGlnArgValGluPheLeuLeuLysAla	2649
D	b	7819	GGCT	CTCTATACGATTCAGTACTCTCTCGGCGACGAGTCGAGTCTCTGGTGAATAACC	7878
Q	y	2650	TrpAla	glusylsAspProMetGlyPheSerTyrAspThrArgCysPheAspSerThr	2669
D	b	7879	TGSA	AATCAAGAAAAACCCCATGGGCTTTATATGACACTCGCTGTGTTCGACTCAACG	7938
Q	y	2670	ValThr	GluArgAspIleArgThrGluLysSerIleTyrArgAlaCysSerLeuProGlu	2689
D	b	7939	GTAC	CCAGAACGACATCCGTGTGTGAGGAGTCAATTTACCNATGTTGACTTGGCCCC	7998
Q	y	2690	GluAla	HisThrAlaIleHisSerLeuThrGluArgLeuTyrValGlyGlyProMetPhe	2709
D	b	7999	GAAG	CCAGACAGCCATATAAATCGTCACAGACGCGCTTTATATCGGGGTCTCTGACT	8058
Q	y	2710	AnS	erLysGlyGlnThrCysGlyTyrArgArgCysArgAlaSerGlyValLeuThrThr	2729
D	b	8059	AAATT	CAAAAGGCGAGAACTCGGTTATCGCGGTGCGCGCGAGCGCGTGTGACGACT	8118
Q	y	2730	SerMet	GlyAasnThrIleThrCysTyrValLysAlaLeuAlaCysLysAlaLaGly	2749
D	b	8119	AGCT	GCGGTAAACCCCTCACATGTTACTTGAAGGCTCTGACGCTTGTGAGCTGCGAAG	8178
Q	y	2750	IleIle	AlaProThrMetLeuValCysGlyAspAspLeuValVallleserGluSerGln	2769
D	b	8179	CTCC	GAGCTGACGATGCTCGTGAAACGAGACGACCTCGTCTGTTATCTGTGAAACGCG	8238
Q	y	2770	GlyThr	GluGluAspGluArgAasnLeuArgAlaPheThrGluAlaMetThrArgTyrSer	2789
D	b	8239	GGAA	CCCAAGGAGGACGCGCGGAGCCCTACGAGTCTTCCAGGAGGCTATGACTAGGTACTCC	8298
Q	y	2790	AlaPro	GlyAspProProArgProGluTyrAspLeuLeuIleThrSerCysSer	2809
D	b	8299	GCCCC	CCCCGGGAGCCGCCCAACCAAGATACGACTTGGAGCTGATAACATCATGTTCC	8358
Q	y	2810	SerA	snValSerValAlaLeuGlyProGlnGlyArgArgTyrTyrLeuThrArgAsp	2829
D	b	8359	TCCA	TGTGTGCGTCCGCCACGATGCATCAGGCAAAAGGGGTGTACTACTCACCCTGAT	8418
Q	y	2830	ProThr	ThrProfileAlaArgAlaLaTrpGluThrValArgHisSerProValAasnSer	2849
D	b	8419	CCCA	CCACCCCTTAGCAGCGGCTGGTGGAGACACTAGACACACTCCAGTTAACTCC	8478

2850 TrpLeuGlyAsnIleGlnTyrAlaProThrIleTyrAlaArgMetValLeuMetThr 2869
 8479 TGGTAGGCAACATTAATGATGATGCGCCCACTTTGTGGCAAGATGATTCGTGACT 8538
 2870 HisPheSerIleLeuMetAlaGlnAspThrLeuAspGlnAsnLeuAsnPheGluMet 2889
 8539 CACTTCTCTCCATCTTCTAGCGCAGGAGCACTTGAAGAGCCCTGGACTGCCAGATC 8598
 2890 TyrGlyAlaValTyrSerValSerProLeuAspLeuProAlaIleIleGluArgLeuHis 2909
 8599 TAGGGGCGCTGTACTTCCATTCAGCCACTTGACCTACCTCAGATCATTTGACGACTCCAT 8658
 2910 GlyLeuAspAlaPheSerLeuHisThrTyrThrProHisGluLeuThrArgValAlaSer 2929
 8659 GGCCTTAGCGCAATTTTCACTCATAGTACTTCCAGGTGAGATCATAGGTGGCTTCA 8718
 2930 AlaLeuArgGlyLeuGlyAlaProProLeuArgAlaTyrLysSerArgAlaValAlaVal 2949
 8719 TGCCTCAGCAAACTTTGGGGTACCACCTTGGAGTCTGGAGACATCGGGCAGGAGCGTC 8778
 2950 ArgAlaSerLeuIleSerArgGlyArgAlaAlaValCysGlyArgTyrLeuPheAsn 2969
 8779 CGCGCTAGCTACTGCTCCAGGAGGAGGGCGCCACTTGTGGCAATACCTCTTCAAC 8838
 2970 TrpAlaValLysThrLysLeuLysLeuThrProLeuProGluAlaArgLeuLeuAspLeu 2989
 8839 TGGCAGTAATAACCAACTTAATCACTCAATCCCGCTGGCTCCCGCTGGACTTG 8898
 2990 SerSerTrpPheThrValGlyAlaGlyGlyAspIleTyrHisSerValSerArgAla 3009
 8899 TCGGCTGTGTTGCTGCTGTACAGCGGGGAGACATATATACAGCCCTGCTCGTGC 8958
 3010 ArgProArgLeuLeuPheGlyLeuLeuLeuLeuPheValGlyValGlyLeuPheLeu 3029
 8959 CGACCCCGTGGTTGCTGCTGCTACTCTACTCTTCTGTAGGGTAGGCATCTACTCTG 9018
 3030 LeuProAlaArg 3033
 9019 CTCCTCCCAACCGA 9030

RESULT 8
 US-09-315-850-13
 Sequence 13, Application US/09315850
 Patent No. 6217872
 GENERAL INFORMATION:
 APPLICANT: OKAYAMA, Hiroto
 APPLICANT: FUKU, Isao
 APPLICANT: MORI, Chisato
 APPLICANT: TAKAMIZAWA, Akahisa
 APPLICANT: YOSHIDA, Iwao
 TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
 TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
 NUMBER OF SEQUENCES: 50
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
 ADDRESSEE: Naughton
 STREET: 1725 K St. N.W. Suite 1000
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20006
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
 SOFTWARE: ASCII
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/315,850
 FILING DATE:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/904,686
 FILING DATE: 01-AUG-1997

APPLICATION NUMBER: US 08/324,977
 FILING DATE: 18-OCT-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 2-167466
 FILING DATE: 25-JUN-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 2-230921
 FILING DATE: 31-AUG-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 2-305605
 FILING DATE: 09-NOV-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/099,706
 FILING DATE: 30-JUL-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/769,996
 FILING DATE: 02-OCT-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/635,451
 FILING DATE: 28-DEC-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: McLeland, Le-Nhung
 REGISTRATION NUMBER: 31,541
 REFERENCE/DOCKET NUMBER: 900703G
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 659-2930
 TELEFAX: (202) 887-0357
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 9030 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA from genomic RNA
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: 1..9030
 OTHER INFORMATION: /note: "sequence = 333 - 9362 of
 OTHER INFORMATION: SEQ ID NO: 1"
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..9030
 US-09-315-850-13
 Alignment Scores:
 Pred. No.: 0 Length: 9030
 Score: 11984.50 Matches: 2175
 Percent Similarity: 83.02% Conservative: 352
 Best Local Similarity: 71.45% Mismatches: 472
 Query Match: 74.35% Indels: 45
 DB: 3 Gaps: 8
 US-09-980-559-2 (1-3033) x US-09-315-850-13 (1-9030)
 QY 1 MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgProGln 20
 Db 1 ATGAGCAGCAATCTTAACCTCAAGAAAAACCAAGTAACACCAACCGCCGCCACAG 60
 QY 21 AspValLysPheProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArg 40
 Db 61 GACGTCAAGTTCCCGGGCGGTGTCAGATCGTTGGTGGAGTTTACCTGTTCGCGCGCAGG 120
 QY 41 GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly 60
 Db 121 GGCCCCAGTTGGGTGTGGCGCGGCCGAGAAAGACTTCCGAGCGGTCCGACACTCTGTGA 180
 QY 61 ArgArgGlnProIleProLysAspArgSerThrGlyLysSerTrpGlyLysProGly 80
 Db 181 AGCGCAGCAACCTATCCCAAGGCTCGCGCGCCGAGGGCAGGACCTGGGCTCAGCCCGGG 240
 QY 81 TyrProTrpProLeuTyrGlyAsnGluGlyLeuGlyTyrAlaGlyTyrLeuLeuSerPro 100
 Db 241 TACCCTTGGCTCTCTATGGCAATGAGGGCTTAGGGTGGGAGGATGGCTCTGTGTACCC 300

```
QY 101 ArgGlySerArgProSerTrpGlyProAsnAspProArgHisArgSerArgAsnValGly 120
Db 301 CGCGGCTCCGGGCTAGTTGGGGCCACAGGACCCCGCGTAGGTGCGTAATTTGGGT 360
QY 121 LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProValVal 140
Db 361 AAGGTCAATGATACCTCATCGCGCTTCGCGGATCTCATGGGGTACATTCGCGTCGTC 420
QY 141 GlyAlaProLeuGlyGlyValAlaArgAlaLeuAlaHisGlyValArgValLeuGluAsp 160
Db 421 GGGGCCCCCTGGGGGGCGCTGCGAGGCGCTGGCACATGGTGTCCGGGTCTGGAGGAC 480
QY 161 GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla 180
Db 481 GCGGTGAATCTATGCAACAGGAATCTGCCGGTTGCTTTTCTATCTTCTCTTGGCT 540
QY 181 LeuLeuSerCysIleThrThrProValSerAlaAlaGluValLysAsnIleSerThrGly 200
Db 541 CTGCTGTCTGCTGACCAACCCAGCTTCGCTTACGAAAGTGCAACGCTGTCGGGATA 600
QY 201 TyrMetValThrAsnAspCysThrAsnAspSerIleThrTrpGlnLeuGlnAlaVal 220
Db 601 TATCATGTCAAGAACGACTGCTTCAACGCAAGCATTTGTGTATGAGCGACGGACTTGATC 660
QY 221 LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTrpIle 240
Db 661 ATGCATACTCTGGGTGCGTCCCTGCGTTCGGGAAGGCACTCTCCGCTGCTGGGTA 720
QY 241 ProValSerProAsnValAlaValGlnArgProGlyAlaLeuThrGlnGlyLeuArgThr 260
Db 721 GCGCTCACTCCCAAGCTCGCAGCAGCAAGCTCACCATCCCAACCAACGACGATACGACG 780
QY 261 HisIleAspMetValMetSerAlaThrLeuCysSerAlaLeuThrValGlyAspLeu 280
Db 781 CACGTGATCTGCTGCTGGGGCGCTGCTTCTGTTCCGCTATGTACGTGGGGGACCTC 840
QY 281 CysGlyGlyValMetLeuAlaGlnMetPheIleValSerProGlnHisIleThrPhe 300
Db 841 TCGGGATCTGTTTCTCTCTCTCTCAGCTGTTTCACTCTCTGCTCGCGGATGTGACA 900
QY 301 ValGlnAspCysAsnCysSerIleThrProGlyThrIleThrGlyHisArgMetAlaTrp 320
Db 901 TTACAGGACTGTAACCTGCTCAATTTATCCGGGCATGTGTGGGTCAACGCTATGGCTGG 960
QY 321 AspMetMetMetAsnTrpSerProThrAlaThrMetIleLeuAlaThrAlaMetArgVal 340
Db 961 GACATGATGATGAACCTGGTGGCCCAACAGCCCTAGTGTGTGCGAGTTACTCCGGATC 1020
QY 341 ProGluValIleIleAspIleIleSerGlyAlaHisTrpGlyValMetPheGlyLeuAla 360
Db 1021 CCACAAGCGCTGCTGGACATGTGGGGGGCCACTGGGGAGTCTCGCGGCTTGGC 1080
QY 361 TyrPheSerMetGlnGlyAlaTrpAlaLysValValIleLeuLeuLeuAlaGly 380
Db 1081 TACTATTCCATGCGGGGAACGGGCTAAGGTTCTGATGTGATGCTACTTTTGTGGC 1140
QY 381 ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr 400
Db 1141 GTTGACGGGGATACCCACGTCAGCGGGGGGGCGCAAGCCAAACCAACAGGCTCGTG 1200
QY 401 SerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTrp 420
Db 1201 TCCATGTTTGCAGTGGGGCGTCTCAGAAATCCAGCTTATTAACCAACCAATGGAGTGG 1260
QY 421 HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer 440
Db 1261 CACATCAACAGGACTGCCCTGAACTGCAATGACTCTCTCCAGACTGGGTTCTTTCGCGG 1320
QY 441 LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg 460
Db 1321 CTGTTCTACACATAGTTTCACTGCTCGGGTGCAGAGCGCATGGCCCGAGTGGCCG 1380
QY 461 SerIleGluAlaPheArgValGlyTrpGlyAlaLeuGlnTyrGluAspAsnValThrAsn 480
Db 1381 ACCAATTGCAAGTTCGACAGGATGGGTGCCATTACTTACTTATGCTGAG-----TCTAGC 1434
QY 481 ProGluAspMetArgProTyrCysTrpHisTyrProArgGlnCysGlyValValSer 500
Db 1435 AGATCAGACACAGAGGCCATATTGCTGGGCACTACCCACCTCCCAATGTACCATCGTACCT 1494
QY 501 AlaLysThrValCysGlyProValTyrCysPheThrProSerProValValValGlyThr 520
Db 1495 GCGTCGGAGTGTGGCGCCAGTACTGCTTACCCCAAGCCCTGCTCGTGGGGGACG 1554
QY 521 ThrAspArgLeuGlyAlaProThrTyrThrTyrGlyGluAsnGluThrAspValPheLeu 540
Db 1555 ACCGATCGTTTCGCTGCTTACGTATAGATGGGGGAGACAGAGACTACGCTGCTGCTG 1614
QY 541 LeuAsnSerThrArgProLeuGlySerTrpPheGlyCysThrTrpMetAsnSerSer 560
Db 1615 CTCAACAACACGCGCGCCGCAAGGCAACTGGTTCGGCTGCACATGGATGAATAGCACC 1674
QY 561 GlyTyrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer 580
Db 1675 GGGTTCACCAACAGATGTGGGGGGCCCCGCTGTAACTCGGGGGGTTCGGCAAC----- 1728
QY 581 ThrAspLeuLeuCysProThrAspCysPheArgLysHisProAspThrThrTyrLeuLys 600
Db 1729 AACACCTTGACCTGCCCCCAGGACTGCTTCCGGAAGCAGCCCCGAGGCTACCTACACAAA 1788
QY 601 CysGlySerGlyProTrpLeuThrProArgCysLeuIleAspTyrProTyrArgLeuTrp 620
Db 1789 TGTGGTTCGGGGCCTTGGCTGACACCTAGTGTGCATGTTGACTATCATACAGGCTCTGG 1848
QY 621 HisTyrProCysThrValAsnThrThrIlePheLysIleArgMetTyrValGlyGlyVal 640
Db 1849 CATTAACCTCGACCTGTTTAACTTTTACCATCTTCAAGTTAGATGATGTGGGGGGGGTG 1908
QY 641 GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp 660
Db 1909 GAGCACAGGCTCAATGCTGCATGCAATTCGACCCGAGGAGAGCGTTGTGACTGGAGGAC 1968
QY 661 ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGluTrpAlaIleLeuPro 680
Db 1969 AGGGATAGCGCGGAGCTCAGCCCGCTGCTGCTCTACACAGAGTGGCAGGTACTGCCCC 2028
QY 681 CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisLeuHisGlnAnile 700
Db 2029 TGTTCTTCCACCCCTACCAGCTCTGTCCACTGTGCTTGATTCACCTCCATCAGAACATC 2088
QY 701 ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrLysTyrIleValArgTrp 720
Db 2089 GTGGAGTGCATACCTATACGTATAGGTCAGCGGTGTCTCTCTTCAATCAATGAG 2148
QY 721 GluTrpValIleLeuLeuPheLeuLeuAlaAspAlaArgValCysAlaCysLeuTrp 740
Db 2149 GAGTATGCTGCTGCTTCTCTCTAGCGGACGACGCTGCTGCTGCTGCTGCTGCTGCTG 2208
QY 741 MetLeuIleLeuLeuGlyGlnAlaGluAlaLeuGluLysValIleLeuHisAla 760
Db 2209 ATGATGTGCTGATACCCAGCGCGCGCTGGAGAACCTGGTGTGCTCTCAATCG 2268
QY 761 AlaSerAlaAlaSerCysAsnGlyPheLeuTyrPheValIlePhePheValAlaIleTrp 780
Db 2269 GGTGTGTGGCGCGCACATGGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2328
QY 781 TyrIleLysGlyArgValValProLeuAlaThrTyrSerLeuThrGlyLeuTrpSerPhe 800
Db 2329 TACATCAAGGAGGCTGCTGCTGCGGCGACATATGCTCTTATGCGGTGCGCGCTG 2388
QY 801 SerLeuLeuLeuAlaLeuProGlnAlaTyrAlaTyrAspAlaSerValHisGly 820
Db 2389 CTCTGCTCTCTGCTGCTTACACCGGAGCTTACCGGAGCTTACCGGAGGAGTGGCTGCA 2448
QY 821 GlnIleGlyAlaAlaLeuValMetIleThrLeuPheThrLeuThrProGlyTyrLys 840
```

Db	2449	TCGTGCGGAGCGCGTCTTTGGTCTGGTACTCTGACCTTGTGACCATACTACAAG	2508	Db	3529	GTGGGCATCTCCGGCTGCTCCGATATGATCACCCCGGGGGTGTGCAAGCGGTGGACTTTGTG	3588
Qy	841	ThrLeuLeuSerArgPheLeuTrpTrpLeuCysTyrLeuLeuThrLeuGlyGluAlaMet	860	Qy	1201	ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAenSerThrPro	1220
Db	2509	GTGTTCTCGCTAGGCTCATATGGTGTACAAATATTTACCAACAGAGCCGAGGCGGAC	2568	Db	3589	CCCGTAGAGTCCATGGAACACTACTATGCGGTCTTCGGTCTTCCAGGACAACACTATCCCC	3648
Qy	861	ValGlnGluTrpAlaProProMetGlnValArgGlyGlyArgAspGlyIleIleTrpAla	880	Qy	1221	ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLys	1240
Db	2569	TTACATGTGGATCCCGCCCTCAACGCTCGGGGAGCGCGATGCCATCATCTCTCTC	2628	Db	3649	CGGCGGTACCCGAGTCATTTCAAGTGCCCACTACACGCTCCCACTGCGAGCGGCAAG	3708
Qy	881	ValAlaIlePheTyrProGlyValValPheAspIleThrIysTrpLeuLeuAlaValLeu	900	Qy	1241	SerThrLysValProValAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAenPro	1260
Db	2629	ATGTGCGCAGTCCATCCAGAGCTAATCTTTGACATCACCAAACTCTTAATTGCCATAC	2688	Db	3709	AGTACTAAAGTCCGCGCTCATATGCAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAG	3768
Qy	901	GlyProAlaTyrLeuLeuLysGlyAlaLeuThrArgValProTyrPheValArgAlaHis	920	Qy	1261	SerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerLysAlaHisGlyLysAenPro	1280
Db	2689	GGTCCGCTCATGTGTCTCAAGCTGCGAGTACCAAGAGTGCCTGCTGCGCGCTCAA	2748	Db	3769	TCGTTGCGCGCTACCTTTAGGGTTTGGGGGTATATGTTCTAAGGCACACCGGTATGAC	3828
Qy	921	AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyGlyArgTyrValGlnMet	940	Qy	1281	AsnIleArgThrGlyValArgThrValThrGlyAlaProIleThrTyrSerThrTyr	1300
Db	2749	GGCTCATTCATGATGCATGTGTAGTGGGAGAGTGCCTGGGGGTCAATTATGTCCAAATG	2808	Db	3829	AACTCAGAACTGGGGTAAGGACCATACACAGGCGCCCGCTCACATCTACTACCTAT	3888
Qy	941	AlaLeuLeuAlaLeuGlyArgTrpThrGlyThrTyrIleTyrAspHisLeuThrProMet	960	Qy	1301	GlyLysPheLeuAlaAspGlyGlyCysAlaGlyAlaTyrAspIleIleLeuLysAsp	1320
Db	2809	GCCTTCATGAAGCTGGCGGCTGCACAGGACGATACATTTACAACTCTTACCCCGCTA	2868	Db	3889	CGCAAGTTTCTGCGCATGCTGCTCTGGGGCGCTTATGACATCATATATATGTAT	3948
Qy	961	SerAspTrpAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIleIlePhe	980	Qy	1321	GluCysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGlnAla	1340
Db	2869	CGGGATTTGGCCACGCGCGGCTTACGAGACCTTGGCGTGGCAGTGGAGCGCGTCTTC	2928	Db	3949	GAGTGCCATTCAACTGACTCGACTCAATCTTGGGCGATCGGCACAGCTCTTGGACCAAG	4008
Qy	981	SerProMetGluLysIleValIleValTrpGlyAlaGluThrAlaAlaCysGlyAspIle	1000	Qy	1341	GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProProGlySerValThr	1360
Db	2929	TCCGACATGAGACCAAGATCATCACTGGGGAGCAGACCCGCGGTGTGGGACATC	2988	Db	4009	GAGACGCTGGAGCGCGGTCTGCTGCTCGCCACCGCTACGCTCCCGGATCGGTCA	4068
Qy	1001	LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp	1020	Qy	1361	ThrProHisProAsnIleGluGluValAlaLeuGlyGlnGluGlyIleProPheTyr	1380
Db	2989	ATCTTGGTCTGCGCTTCCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	3048	Db	4069	GTGCCACACCAAAATCATCGAGAGTGGCGCTCTTAATCTGAGAGATCTCCCTTCTAT	4128
Qy	1021	GlyTyrThrSerLysGlyTrpSerLeuLeuAlaProIleThrAlaTyrAlaGlnThr	1040	Qy	1381	GlyArgAlaIleProLeuSerTyrIleLysGlyGlyArgHisLeuIlePheCysHisSer	1400
Db	3049	AGTCTTGAAGGCGGGGTTGCGACTCTCGCGCCCATCAACGCTTCTCCCAACAGACG	3108	Db	4129	GGCAAGCCATCCCATTAAGCCATCAGGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAG	4188
Qy	1041	ArgGlyLeuLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGluGlnAla	1060	Qy	1401	LysLysLysCysAspGluLeuAlaAlaLeuArgGlyMetGlyLeuLeuAenSerValAla	1420
Db	3109	CGGGGCTCTTGTGTGATCATCATCATAGCTTACAGCGCGGAGCAAGAACAGGTGCGAG	3168	Db	4189	AGAAGAAGTGGCAGAGCTCCCGCAAGCTGTCAAGCTCGGAATCAAGCTGTGGCG	4248
Qy	1061	GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly	1080	Qy	1421	TyrTyrArgGlyLeuLeuAspValSerValIleProThrGlnGlyAspValValValAla	1440
Db	3169	GGAGAGTTTCAAGTGGTTTCCACCGCAACACATCTCTTCTGCGGACCTCGGTCAACGC	3228	Db	4249	TATTACCGGGGCTCGATGTCTCGTATACCAACTATCGGAGACGCTGTGTGTGGCA	4308
Qy	1081	ValLeuTrpThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro	1100	Qy	1441	ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAenVal	1460
Db	3229	GTGTGTGGACCGTTTACCATGTGTGCTGCTCAAGACCTTAGCGCGCGCAAGGGGCCA	3288	Db	4309	ACAGACGCTCTGATGACGGGCTATACGGGCGACTTTGACTCAGTGTGACTGTATACACA	4368
Qy	1101	ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTrpProSerProGly	1120	Qy	1461	AlaValThrGlnValValAspPheSerLeuAspProThrPheThrIleThrGlnIle	1480
Db	3289	ATCACCAGATGTACACTAATGTGGACCAAGGACCTGTGCGCTGGCGCAAGCCCGCGG	3348	Db	4369	TGTGTCCACGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4428
Qy	1121	ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArgAsn	1140	Qy	1481	ValProGlnAspAlaValSerArgSerGlnArgArgGlyArgThrGlyArgGlyArgLeu	1500
Db	3349	CGCGGTTCTTGACACCATGACCTGTGGGAGCTGACAGCTTACTTGTGTACAGACAT	3408	Db	4429	GTGCTTCAAGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4488
Qy	1141	AlaAspValIleProAlaArgArgArgGlyAspLysArgGlyAlaLeuLeuSerProArg	1160	Qy	1501	GlyIleTyrArgTyrValSerThrGlyGluArgAlaSerGlyMetPheAspSerValVal	1520
Db	3409	GCTGACGCTATCCGGTGGCGCGCGCGCGCGAGTACAGTGGGAGGAGGAGGAGGAGGAG	3468	Db	4489	GGCATCTACAGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4548
Qy	1161	ProLeuSerThrLeuLysGlySerSerGlyProValLeuCysProArgGlyHisAla	1180	Qy	1521	LeuCysGluCysTyrAspAlaGlyAlaAlaTyrTyrGluLeuThrProSerGluThrThr	1540
Db	3469	CCTGTCTCTTACTTGAAGGCTCTTCTGGGTGGTCCACTGCTCTGCTGCTGCTGCTGCT	3528	Db	4549	CTGTGTGAGTGTATACCGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	4608
Qy	1181	ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle	1200	Qy	1541	ValArgLeuArgAlaTyrPheAenThrProGlyLeuProValCysGlnAspHisLeuGlu	1560
				Db	4609	GTGTGCTGCGGCTTACCTGAACACACAGGCTTGGCGGTTGGCGGAGGAGGAGGAGGAG	4668

```
QY 1561 pheTrpGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThr 1580
Db 4669 TTCTGGAGAGTGTCTTACAGCGCTCACCCATATAGATGACACTCTTGTGCCAGACC 4728
QY 1581 LysGlnSerGlyGluAsnPheAlaTyrLeuThrAlaTyrGlnAlaThrValCysAlaArg 1600
Db 4729 AAGCAGCAGGAGCAACATTCCTACCTGGTAGCATACCAAGCCACCGGTGTGCGCCAGG 4788
QY 1601 AlaLysAlaProProSerTrpAspValMetTrpLysCysLeuThrArgLeuLysPro 1620
Db 4789 GCTCAGGCCCCCTCCATCATGGGATCAATGTGAAGTGTCTCATAGCGCTGAACCT 4848
QY 1621 ThrLeuValGlyProThrProLeuLeuTyrArgLeuGlySerValThrAsnGluValThr 1640
Db 4849 AGCTGCACGGGCCAACCCCTTGTCTGACAGCTCGGAGCGCTCCAGAATGAGGTCCACC 4908
QY 1641 LeuThrHisProValThrLysTyrIleAlaThrCysMetGlnAlaAspLeuGluValMet 1660
Db 4909 CTCACCCACCCCAATAACCAATACATCATGGCATGATGTGGCTGACCTGGAGGTGTC 4968
QY 1661 ThrSerThrTrpValLeuAlaGlyGlyValLeuAlaValAlaAlaTyrCysLeuAla 1680
Db 4969 ACTAGACCTTGGGTGTGGTGGCGGAGTCTTTCAGCTCTGGCCCGGTATTGCCCTGACA 5028
QY 1681 ThrGlyCysValCysIleIleGlyArgLeuHisIleAsnGlnArgAlaValValAlaPro 1700
Db 5029 ACAGGCAGTGTGTCTATTGTGGTAGAATATCTTGTCCGGGAGCGCGCCATTGTCTCC 5088
QY 1701 AspLysGluValLeuTyrGluAlaPheAspGluMetGluGluCysAlaSerArgAlaAla 1720
Db 5089 GACAGGAGCTTCTCTACAGGAGTTCGATGAAATGGAAGAGTGGCCCTCGCACCTCCCT 5148
QY 1721 LeuIleGluGlyGlnArgIleAlaGluMetLeuLysSerLysIleGlnGlyLeuLeu 1740
Db 5149 TATCATCGACGAGGAATGAGTCTCGCCGAGCAATTCAAGCAGAAAGCGCTCGGGTTACTG 5208
QY 1741 GlnGlnAlaSerLysGlnAlaGlnAspIleGlnProThrValGlnAlaSerTrpProLys 1760
Db 5209 CAACACAGCCACCAACAGCGGAGGTGTGTCTCCGTGGAGTCCAAGTGGCGAGCC 5268
QY 1761 ValGluGlnPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeuAla 1780
Db 5269 CTTGAGACATCTTGGCGGAGCAGACATGTGGAATTTTCATCAGCGGATACAGTACTTAGCA 5328
QY 1781 GlyLeuSerThrLeuProGlyAsnProAlaValAlaSerMetMetAlaPheSerAlaAla 1800
Db 5329 GGCCTTATCCACTCTGCTGGGAACCCCGCAATAGCATCATTTGATGGCATTTACAGCCTCT 5388
QY 1801 LeuThrSerProLeuSerThrThrThrIleLeuLeuAsnIleLeuGlyGlyTrpLeu 1820
Db 5389 ATCACCAGCCCGCTCACCCACCAAGTACCTCTGTTAAACATCTTGGGGGGTGGTG 5448
QY 1821 AlaSerGlnIleAlaProAlaGlyAlaThrGlyPheValValSerGlyLeuValGly 1840
Db 5449 GCTGCCAACTGCCCCCGCCAGCGCGCTTGGCTTTCTGGGCGCCGCGCATCGCGGT 5508
QY 1841 AlaAlaValGlySerIleGlyLysValLeuValAspIleLeuAlaGlyTyrGly 1860
Db 5509 GCGGCTGTGGTAGCATAGCCCTTGGGAAGGTGCTTGTGGACATCTTGGCGGGTTATGA 5568
QY 1861 AlaGlyIleSerGlyAlaLeuValAlaPheLysIleMetSerGlyGluLysProSerMet 1880
Db 5569 GCAGGAGTGGCCGCGCTGTGGCTTTAGGTCTATGAGCGCGAGATGCGCTCCACC 5628
QY 1881 GluAspValValAsnLeuLeuProGlyIleLeuSerProGlyAlaLeuValValGlyVal 1900
Db 5629 GAGGACCTGTGCTCAATCTACTTCTGCTGCTCTCTCTCTGCGGCCCTGTGTCGGGTC 5688
QY 1901 IleCysAlaAlaIleLeuArgArgHisValGlyProGlyGluAlaValGlnTrpMet 1920
Db 5689 GTGTGTGCAACAACTACTGCTGCACAGTGGGTCCGGGAGAGGGGGCTGTGCAGTGGATG 5748
QY 1921 AsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaProThrHisTyrValThr 1940
Db 5749 AACCGCTGATAGCTTGGCTCGCGGGTAAATCATGTTTCCCAACGCACTATGTGCT 5808
QY 1941 GluSerAspAlaSerGlnArgValThrGlnLeuLeuGlySerLeuThrIleThrSerLeu 1960
Db 5809 GAGAGCGAGCGCCAGCGCTTACTCAGATCTCTCAGGCTTACCATCATCTACGCTG 5868
QY 1961 LeuArgGlyLeuHisAsnTrpIleThrGluAspCysProIleProCysGlyGlySerTrp 1980
Db 5869 CTGAAAAGGCTCCACAGTGGATTAAATGAAGACTGCTCCACACCGTGTCCGGCTGCTG 5928
QY 1981 LeuArgAspValTrpAspTrpValCysThrIleLeuThrAspPheLysAsnTrpLeuThr 2000
Db 5929 CTAAGGATGTTTGGGACTGGATATGACAGGTGTTGACTGACTCAAGACCTGGCTCCAG 5988
QY 2001 SerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTyrLys 2020
Db 5989 TCCAAGCTCTCTCCCGCAGCTACTCGAGTCCCTTTTCTCGTCCACACGCGGTACAAG 6048
QY 2021 GlyValTrpAlaGlyThrGlyIleMetThrThrArgCysProCysGlyAlaAsnIleSer 2040
Db 6049 GGAGTCTGGCGGGAGAGCGGCATCATGCAACCACTTGCCCATGTGGAGCACAGATCACC 6108
QY 2041 GlyAsnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTrp 2060
Db 6109 GCAGATGTCAAAAACGTTTCCATGAGTATCGTGGGCTTAAGACCTGCAGCAACACGTGG 6168
QY 2061 GlnGlyThrPheProIleAsnCysTyrThrGluGlyGlnCysValProLysProAlaPro 2080
Db 6169 CATGGAACATTTCCCATCAACGATACACACGCGGCCCTCGCACACCTCTCCAGCGCCA 6228
QY 2081 AsnPheLysValAlaIleTrpArgValAlaAlaSerGluTyrAlaGluValThrGlnHis 2100
Db 6229 AACTATTCTAGGCGCTGTGGCGGTGGCGCTGAGGAGTCTGAGGAGTCAAGCGGTG 6288
QY 2101 GlySerTyrHisTyrIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeu 2120
Db 6289 GGGGATTTCCACTACGTGACGGGCATGACCACTGACCAACGTAAGTGCCCATGCCAGTT 6348
QY 2121 ProSerProGluPhePheSerTrpValAspGlyValGlnIleHisArgPheAlaProThr 2140
Db 6349 CCGGCTCTCCGAATTTCTTCGGAGGTGGACGAGTGGCGGTTCACAGGTACGCTCCGGCG 6408
QY 2141 ProLysProPhePheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal 2160
Db 6409 TCAGCGCTCTCTCCTAGGAGGAGGTATCTTCCAGTCCGGCTCAACCAATACCTGGTT 6468
QY 2161 GlySerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThr 2180
Db 6469 GGTTCACAGCTACCATGCGAGCCCGAACCGGATGTAGCAGTCTCACTTCCATGCTCACC 6528
QY 2181 AspProSerHisIleThrAlaGluThrAlaAlaArgArgLeuAlaArgGlySerProPro 2200
Db 6529 GACCCCTCCCATCACAGCAACCGCTAAGCGGTAGGTGGCCAGGGGGTCTCCCCCG 6588
QY 2201 SerGluAlaSerSerAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThr 2220
Db 6589 TCCTTGGCCAGCTCTTCTAGCTAGCCAGTGTCTTGGCCCTTCTTGAAGCGCATGCACT 6648
QY 2221 ThrHisGlyLysAlaTyrAspValAspMetValAspAlaAsnLeuPhe 2236
Db 6649 ACCCACTGCTCTCTCCGAGCGCTGACCTCATCGAGGCCCACTCTCTGTGGCGGAGGAG 6708
QY 2237 MetGlyGlyAspValThrArgIleGluSerGlySerLysValValValLeuAspSerLeu 2256
Db 6709 ATGGCGCGGAGCAATCACCGCGTGGAGTCGGAGAACAGGTGTGTCTCTGGACTCTTTC 6768
QY 2257 AspProMetValGluGluArgSerAspLeuGluProSerIleProSerGluTyrMetLeu 2276
Db 6769 GACCCGCTTTCGAGCGGAGGAGGATGAGAGGAGTATCCGTTCCGCGGAGATCTTCGCGG 6828
QY 2277 ProLysLysArgPheProAlaLeuProAlaTrpAlaArgProAspTyrAsnProPro 2296
```


Db	6829	AAATCCAGAAAGTTCCCGCAGCATGCCATCTGGCGCGCCGGATTACAAACCTCCA	6888	7819	GGCTCTCATACGAGTCCAGTACTCTCTGGCGCAGCGAGTGGAGTCTCTGGTGAATACC	7878
Qy	2297	LeuValGluSerTrpLysArgProAspTyrGlnProAlaThrValAlaGlyCysAlaLeu	2316	2650	TrpAlaGluLysLysLeuAspProMetGlyPheSerTyrAspThrArgCysPheAspSerThr	2669
Db	6889	CTGTTAGAGTCTCGGAAGACCGCGACTACGTCCTCCCGTGGTGCACCGGGTGG	6948	7879	TGGAATCAAGAAAACCCCATGGCTTTTCATATGACACTCGCTGTTTCGACTCAAG	7938
Qy	2317	ProProArgLysThrProThrProProArgArgArgThrValGlyLeuSer	2336	2670	ValThrGluArgAspIleArgThrGluGluSerIleTyrArgAlaCysSerLeuProGlu	2689
Db	6949	CAACCTATCAGGCCCTCAATACCACTCCACGAGAAAGAGCGGTGTCCTAAC	7008	7939	GTCCCGGAGAACGACATCCGTTGAGGAGTCAATTTACCAATGTTGACTTGGCCCC	7998
Qy	2337	GluAspSerIleGlyAspAlaLeuGlnLeuAlaIleLysSerPheGlyGlnProPro	2356	2690	GluAlaHisThrAlaIleHisSerLeuThrGluArgLeuTyrValGlyGlyProMetPhe	2709
Db	7009	GAGTCTCCGTGCTCTGCTTACGAGGAGTCTTACTTAAGACCTTCGGC	7059	7999	GAAGCCAGACAGCCATAAATCGCTCAGAGCGGCTTATATCGGGGCTCTCTGACT	8058
Qy	2357	ProSerGlyAspSerGlyLeuSerThrGlyAlaGlyAlaAlaAspSerGlySerGlnThr	2376	2710	AsnSerLysGlyGlnThrCysGlyTyrArgArgCysArgAlaSerGlyValLeuThrThr	2729
Db	7060	-----AGTCCGAATCATCGCGCTCGACAGCGCGACGAC	7098	8059	AATTCAAGGGCAGAACTGCGGTATCGCGGTCCGCGCGAGCGGCTGTGAGACT	8118
Qy	2377	Pro---ProAspGluLeuAla-----LeuSerGluThrGlySerIleSer	2390	2730	SerMetGlyAsnThrIleThrCysTyrValLysAlaLeuAlaAlaCysLysAlaGly	2749
Db	7099	GCCCTTCTGACAGGCTCCGACGAGCGGTGCAAAAGGATCCGACGTTGAGTCTGACTCC	7158	8119	AGTCGGGTAAACCCCTCACATGTTACTTTGAAGGCTCTGCGAGCTGTGAGCTCGAAG	8178
Qy	2391	SerMetProProLeuGluGlyGluLeuGlyAspProAspLeuGluProGluGlnValGlu	2410	2750	IleLeuAlaProThrMetLeuValCysGlyAspAspLeuValIleSerGluSerGln	2769
Db	7159	TCCATGCCCTTGGGGGGAACCGGGGACCCCGAATCTC-----	7200	8179	CTCCAGGACTGACGATGCTCGTGAACGAGACGACCTCGTCTGTTATCTGTGAAAGCGCG	8238
Qy	2411	ProGlnProProGlnGlyValAlaAlaProGlySerAspSerGlySerTrpSer	2430	2770	GlyThrGluGluAspGluArgAsnLeuArgAlaPheThrGluAlaMetThrArgTyrSer	2789
Db	7201	-----AGTACGGGCTTGTGTCT	7218	8239	GAAGCCCAAGAGGACGCGGAGCTTACGAGTCTTACGAGGAGCTATGACTAGTACTCC	8298
Qy	2431	ThrCysSerGluGlu---AspAspSerValValCysSerMetSerTyrSerTrpThr	2449	2790	AlaProProGlyAspProProArgProGluTyrAspLeuGluLeuIleThrSerCysSer	2809
Db	7219	ACCGTACGAGGAGGAGTGTAGTGGAGTGTCTGCTGCTCAATGTCTACATGAGCA	7278	8299	GGCCCCCGGGGACCGCCCCCAACCAAGATACGACTTGGAGCTGATAACATCATGTTCC	8358
Qy	2450	GlyAlaLeuIleThrProCysSerProGluGluGluLysLeuProIleAsnProLeuSer	2469	2810	SerAsnValSerValAlaLeuGlyProGlnGlyArgArgTyrTyrLeuThrArgAsp	2829
Db	7279	GGCGCTTGATACGCCATCGCTGCGGAGGAAGCAAGCTGCCATCAACGGTTGAGC	7338	8359	TCCAATGTGCTGCTGCCACGATGCATCAGGCAAAAGGGTGTACTACCTCCCGGTGAT	8418
Qy	2470	AsnSerLeuLeuArgTyrHisAsnLysValTyrCysThrThrThrLysSerAlaSerLeu	2489	2830	ProThrThrProIleAlaArgAlaAlaTrpGluThrValArgHisSerProValAsnSer	2849
Db	7339	AACTCTTGTGCGCCACCAATACATGTTTATGCCACAACATCTCGCAGCGCAGCCTG	7398	8419	CCACACCCCTTAGCAGGCTGCTGGGAGACAGCTAGACACACTCCAGTTACTCC	8478
Qy	2490	ArgAlaLysLysValThrPheAspArgMetGlnValLeuAspSerTyrTyrAspSerVal	2509	2850	TrpLeuGlyAsnIleIleGlnTyrAlaProThrIleTrpAlaArgMetValLeuMetThr	2869
Db	7399	CGCAGAGAGAGTACCTTTGACACTGCACTGCACTGCACTGCACTACCGGACGCTG	7458	8479	TGCTAGGCAACATTTATGATGATGCGCCCTTGTGGCAAGGATGATCTGTGACT	8538
Qy	2510	LeuLysAspIleLysLeuAlaAlaSerLysValThrAlaArgLeuLeuMetGluGlu	2529	2870	HisPhePheSerIleLeuMetAlaGlnAspThrLeuAspGlnAsnLeuAsnPheGluMet	2889
Db	7459	CTCAGGAGATGAAGGAGGCGTCCACAGTTAAGGCTAAACTCTCTATCCGTAGAGAA	7518	8539	CACCTTCTCTCCATCTTAGCGGAGGACCACTTGAAGGCTGAGCTGCCAGATC	8598
Qy	2530	AlaCysGlnLeuThrProProHisSerAlaArgSerLysTyrGlyPheGlyAlaLysGlu	2549	2890	TyrGlyAlaValTyrSerValSerProLeuAspLeuProAlaIleIleGluArgLeuHis	2909
Db	7519	GCCTGCAAGTGCAGCCCCACATTCGCGCAATCCAAAGTTTGGCTATGGGCAAGGAC	7578	8599	TACGGGCGCTGTACTTCCATGAGGCACTTGACCTACCTCAGATCAATGACACCTCCAT	8658
Qy	2550	ValArgSerLeuSerGlyArgAlaValAsnHisIleLysSerValTrpLysAspLeuLeu	2569	2910	GlyLeuAspAlaPheSerLeuHisThrTyrThrProHisGluLeuThrArgValAlaSer	2929
Db	7579	GTCCGAACCTATCCAGAGGCGGTAAACCATCCACTCCCTGTGGAGAGGACTTGTCTG	7638	8659	GGCCTTAGGCACTTTTCACTCCATGATGATCTCCAGGTGAGATCAATAGGTGGTCTCA	8718
Qy	2570	GluAspSerGluThrProIleProThrIleMetAlaLysAsnGluValPheCysVal	2589	2930	AlaLeuArgLysLeuGlyAlaProProLeuArgAlaTrpLysSerArgAlaAlaVal	2949
Db	7639	GAAGACACTGTGACACCAATGACACCATCATGCGCAAAATGAGGTTTCTGTGTC	7698	8719	TGCCTCAGGAAACTTGGGTGAGGAGGCGGCGCTTGGGCAAAATACCTCTTCAAC	8778
Qy	2590	AspProThrLysGlyLysLysAlaAlaArgLeuIleValTyrProAspLeuGlyVal	2609	2950	ArgAlaSerLeuIleSerArgGlyArgAlaAlaValCysArgValTyrLeuPheAsn	2969
Db	7699	CAACAGAGAAAGGAGCGCTTAAGCCAGCGGCTTATCGTATTTCCAGATCTCGGAGTC	7758	8779	CGCGCTAGGCTACTGTGCCAGGAGGAGGCGGCGCTTGGGCAAAATACCTCTTCAAC	8838
Qy	2610	ArgValCysGlnLysMetAlaLeuTyrAspIleThrGlnLysLeuProGlnAlaValMet	2629	2970	TrpAlaValLysThrLysLeuLysLeuThrProLeuProGluAlaAlaArgLeuLeuAspLeu	2989
Db	7759	CGTGATGCGAGAGATGGCCCTCTATGATGTGCTCTCCACCTTCTCAGGTGCGGATG	7818	8839	TGGGAGTAAACCAACTTAACCTCACTCCAAATCCCGGCTGCGTCCCGGCTGGACTTG	8898
Qy	2630	GlyAlaSerTyrGlyPheGlnTyrSerProAlaGlnArgValGluPheLeuLysAla	2649	2990	SerSerTrpPheThrValGlyAlaGlyGlyArgPheIleTyrHisSerValSerArgAla	3009
Db				8899	TCCGGCTGCTGCTGCTGCTTACAGCGGGGAGACATATATACAGCCCTCTCTCGTCC	8958

QY 3010 ArgProArgLeuLeuPheGlyLeuLeuLeuPheValGlyValGlyLeuPheLeu 3029
 Db 8959 CGACCCCGTGGTTTCATGCTGCTACTCTCTACTTTCTGTAGGGTAGGCATCTACCTG 9018

QY 3030 LeuProAlaArg 3033
 Db 9019 CTCCCCAACCGA 9030

RESULT 9

US-08-324-977-1
 ; Sequence 1, Application US/08324977
 ; Patent No. 5747339
 ; GENERAL INFORMATION:
 ; APPLICANT: OKAYAMA, Hiroto
 ; APPLICANT: FUKU, Isao
 ; APPLICANT: MORI, Chisato
 ; APPLICANT: TAKAMIZAWA, Akahisa
 ; APPLICANT: YOSHIDA, Iwao
 ; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
 ; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
 ; NUMBER OF SEQUENCES: 50
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Armstrong, Westerman, Hattori, Mclelland &
 ; ADDRESSEE: Naughton
 ; STREET: 1725 K St. N.W. Suite 1000
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20006
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/324,977
 ; FILING DATE: 18-OCT-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 2-167466
 ; FILING DATE: 25-JUN-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 2-230921
 ; FILING DATE: 31-AUG-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 2-305605
 ; FILING DATE: 09-NOV-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/099,706
 ; FILING DATE: 30-JUL-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/769,996
 ; FILING DATE: 02-OCT-1991
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/635,451
 ; FILING DATE: 28-DEC-1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Stevens-Smith, Theresa M.
 ; REGISTRATION NUMBER: 36,281
 ; REFERENCE/DOCKET NUMBER: 900703D
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 659-2930
 ; TELEFAX: (202) 887-0357
 ; TELEX: 440142
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 9416 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cdna from genomic RNA
 ; FEATURE:
 ; NAME/KEY: CDS

LOCATION: 333...9362

US-08-324-977-1

Alignment Scores:

Pred. No.: 0 Length: 9416
 Score: 11984.50 Matches: 2175
 Percent Similarity: 83.02% Conservative: 352
 Best Local Similarity: 71.45% Mismatches: 472
 Query Match: 74.35% Indels: 45
 DB: 1 Gaps: 8

US-09-980-559-2 (1-3033) x US-08-324-977-1 (1-9416)

QY 1 MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGln 20
 Db 333 ATGAGCAGCAATCTTAAACCTCAAAGAAAAACCAACGTAACACCAACCGCCACAG 392
 QY 21 AspValLysPheProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArg 40
 Db 393 GACGTCAAGTTCCCGCGCGGTGCTCAGATCGTTGGTGGAGTTTACCTGTTCGCGCAGG 452
 QY 41 GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly 60
 Db 453 GGCCCCAGGTGGGTGGCGCGCCGAGGAAGACTTCCGAGCGGTCCGAACCTCTGGGA 512
 QY 61 ArgArgGlnProIleProLysAspArgArgSerThrGlyLysSerThrGlyLysProGly 80
 Db 513 AGCGCAGCAACCTATCCCCAAGGCTCGCGCGCCGAGGCGAGGACCTGGGCTCAGCCCGG 572
 QY 81 TyrProTrpProLeuTyrGlyAsnGluGlyLeuGlyTyrAlaGlyTyrLeuLeuSerPro 100
 Db 573 TACCCCTTGGCCCTCTCTATGGCAATGAGGGCTTATGGGTGGCGAGGATGGCTCTGTCA 632
 QY 101 ArgGlySerArgProSerTrpGlyProAsnAspProArgHisArgSerArgAsnValGly 120
 Db 633 CGCGGCTCCCGCCCTAGTTGGGGCCCCACGAGGACCCCGCGGTAGGTGGCGTAATTGGGT 692
 QY 121 LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProValVal 140
 Db 693 AAGGTCATCGATACCTCATATGCGGCTTCGCGCGATCTCATGGGGTACATTCGCTCGTC 752
 QY 141 GlyAlaProLeuGlyGlyValAlaAlaAlaGluAlaHisGlyValArgValLeuGluAsp 160
 Db 753 GGCGCCCCCTGGGGGGCGCTGCCAGGCGCTGGCAGCATGGTGTCCGGGTTCGAGGAGAC 812
 QY 161 GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla 180
 Db 813 GCGGTGAACATATGCAACAGGGAATCTGCCCGGTGCTCTTTTCATCTTCCTCTTGGCT 872
 QY 181 LeuLeuSerCysIleThrThrProValSerAlaAlaGluValLysAsnIleSerThrGly 200
 Db 873 CTGCTGTCTGCTGCTGACACCCAGCTTCGGCTTACGAAGTGCACACGCTGTCGGGATA 932
 QY 201 TyrMetValThrAsnAspCysThrAsnAspSerIleThrTrpGlnLeuGlnAlaVal 220
 Db 933 TATCATGTCAAGAACGACTGCTCCAAACGCAAGCATTTGTATGAGGACGAGGACTTGATC 992
 QY 221 LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTrpIle 240
 Db 993 ATGCATATCTCTGGGTGGCGTGGCTGCGGTTCGGGAAGGCAACTCTCTCCGCTCGGTGA 1052
 QY 241 ProValSerProAsnValAlaValGlnArgProGlyAlaLeuThrGlnGlyLeuArgThr 260
 Db 1053 GGGCTCACTCCACGCTCGAGCCAGGAAGCTCACCATCCCAACGACGATACGACGC 1112
 QY 261 HisIleAspMetValValMetSerAlaThrLeuCysSerAlaLeuTyrValGlyAspLeu 280
 Db 1113 CACGTGATCTGCTGTTGGGGCGGTGCTTTCTGTTCGCTATGTACGTGGGGGACCTC 1172
 QY 281 CysGlyGlyValMetLeuAlaAlaGlnMetPheIleValSerProGlnHisHisTrpPhe 300
 Db 1173 TCGGATCTGTTTCTCTCTCTCAGCTGTTCACCTTCTCGCTCGCGCATGTGACA 1232

Copied from 09540843 on 05/19/2004

Db 3381 AGTCTTGAAGGGCGGGGTTGGACATCTCTCGCCGCCCATCACGGCCCTACTCCCAACAGACG 3440
Qy ArgGlyLeuLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGluGlnAla 1060
Db CGGGGCTACTTGGTTGCATCATCACTAGCTTTACAGGCGCGGCAAGAACACAGGTCTGAG 3500
Qy GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly 1080
Db CGAGAGGTTTCAGGTGGTTCACCGCAACACAAATCTCTCTGGCGAGCTCGGTCAACGGC 3560
Qy ValLeuTrpThrValThrHisGlyAlaGlyAsnLysThrIleAlaGlySerArgGlyPro 1100
Db GTGTGTGGACCGTTTACCATGTGTCTGGCTCAAGACCTTAGCCGCGCAAGGGGGCCA 3620
Qy ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTrpProSerProGly 1120
Db ATCACCCAGATGTACACTAATGTGGACAGGACCTCTGTCGGCTGGCCCAAGCCCCCGG 3680
Qy ThrLysSerLeuGlnProCysThrCysGlyAlaValAspLeuTyrLeuValThrArgAsn 1140
Db GCGCGTTCCTTGCACACCATGCACTGTGGCAGCTCAGACCTTTACTTGTGTACAGACAT 3740
Qy AlaAspValIleProAlaArgArgGlyAspLysArgGlyValAlaLeuLeuSerProArg 1160
Db GCTGACGTCAATCCGGTGGCGCGCGGGGCGACAGTAGGGGGAGCTGTCTCTCCCCAG 3800
Qy ProLeuSerThrLeuLysGlySerSerGlyGlyProValLeuCysProArgGlyHisAla 1180
Db CCTGTCTCTACTTGAAGGGCTCTTCGGGTGGTCTCACTGCTCTGCCCTTCGGGCACGCT 3860
Qy ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle 1200
Db GTGGGATCTTCGGGCTGCGGTATGACACCGGGGGGTTCGGAAGCGGTGGACTTTGTG 3920
Qy ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro 1220
Db CCGGTAGATGCTCATGAACTACTATGCGGTCTCCGGTCTTCACGGACAACCTCATCCCC 3980
Qy ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLys 1240
Db CGGGCGTACCGGAGTCAATTCAGTGGCCACCTACAGCTCCCACTGGCAGGGCAAG 4040
Qy SerThrLysValProValAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnPro 1260
Db AGTACTAAAGTGCAGTGCATATGACAGCCCAAGGTACAAGGTGCTGCTCTCAATCG 4100
Qy SerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerLysAlaHisGlyIleAsnPro 1280
Db TCCGTTGGCGCTACCTTAGGGTTTGGGGGCTATATGCTTAAGGCACACGGTATTGACCC 4160
Qy AsnIleArgThrGlyValArgThrValThrThrGlyAlaProIleThrTyrSerThrTyr 1300
Db AACATCAGAACTGGGGTAGGACCATACACAGGGCGCCCCGTACATACCTCTACTAT 4220
Qy GlyLysPheLeuAlaAspGlyGlyCysAlaGlyGlyAlaTyrAspIleIleCysAsp 1320
Db GGCAGTTCCTTGGCCGATGGTGTCTCTGGGGGCGCTTATGACATCATATATGTAT 4280
Qy GluCysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGlnAla 1340
Db GAGTGCATTAACATCGATCGATCAATCTTTGGGCATCGGCACAGCTCTCGACCAAGCG 4340
Qy GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProGlySerValThr 1360
Db GAGACGGCTGGAGCGCGCTTGTGCTGCTCGCCACCGCTACCGCTCCGGGATCGGTACC 4400
Qy ThrProHisProAsnIleGluGluValAlaLeuGlyGlnGluGlyGluIleProPheTyr 1380
Db GTGCCACACCAACATCAGAGAGGTGGCCCTGTCTAATCTGGAGAGATCCCTCTAT 4460
Qy GlyArgAlaIleProLeuSerTyrIleLysGlyArgHisLeuIlePheCysHisSer 1400
Db GGCNAAGCCATCCCATTTGAAGCCATCAGGGGGGGAAGGCATCTCATTTTCTGTCTCC 4461
Db

Qy LysLysLysCysAspGluLeuAlaAlaLeuArgGlyMetGlyLeuAsnSerValAla 1420
Db AAGAAGAAGTGCAGAGCTCGCCCAAGCTGTCCAGGCTCGAATCAACCTGTGGG 4580
Qy TyrTyrArgGlyLeuAspValSerValIleProThrGlnGlyAspValValValAla 1440
Db TATTACCGGGGCTCGATGTGTCCGTATACCACTATTCGAGAGCTGTGTGTGTGGCA 4640
Qy ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnVal 1460
Db ACAGACGCTCTCATGACGGCTATACGGCGACTTTGACTCAGTGATCGACTGTATACACA 4700
Qy AlaValThrGlnValValAspPheSerLeuAspProThrPheThrIleThrGlnIle 1480
Db TGTGTCAACCCAGACAGTCAGCTTCAGTTGGATCCACCTTCACCATTCGAGACGACGACC 4760
Qy ValProGlnAspAlaValSerArgSerGlnArgArgGlyArgThrGlyArgGlyArgLeu 1500
Db GTGCCCTCAAGACGAGTGTCCGCTCGCAGCGCGGGGTAGGACTGGCAAGGGGTAGGAGA 4820
Qy GlyIleTyrArgTyrValSerThrGlyGluArgAlaSerGlyMetPheAspSerValVal 1520
Db GGCATCTACAGGTTCGTGACTCCGGAGAACGGCCCTCGGGATGTTTCATTCCTCGGTC 4880
Qy LeuCysGluCysTyrAspAlaGlyAlaAlaTyrTyrGluLeuThrProSerGluThrThr 1540
Db CTGTGTGAGTGTATGACGCGGGCTGTGCTGTGTACGAGCTCACCCCGCCGAGACTCG 4940
Qy ValArgLeuArgAlaTyrPheAsnThrProGlyLeuProValCysGlnAspHisLeuGlu 1560
Db GTTAGGTTCGGGGCTTACCTGAACACACACAGGGTTCGGCTTTCGAGACCACTCGGAG 5000
Qy PheTrpGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThr 1580
Db TTCTGGGAGAGTGTCTTACAGGCTTACCCCATATAGATGACACTTCTTGTCCACAGCC 5060
Qy LysGlnSerGlyGluAsnPheAlaTyrLeuThrAlaTyrGlnAlaThrValCysAlaArg 1600
Db AAGCAGGCGAGGAGACACTTCCCTTACCTGTGTAGTACCAAGCCACGGTGTGGCCAGG 5120
Qy AlaLysAlaProProSerTrpAspValMetTrpLysCysLeuThrArgLeuLysPro 1620
Db GCTCAGGCCCCACCTCCATCATGGGATCAAAATGTGGAAGTGTCTCATACGGCTGAAACCT 5180
Qy ThrLeuValGlyProThrProLeuLeuTyrArgLeuGlySerValThrAsnGluValThr 1640
Db ACGCTGCACGGGCCCAACCTTGTCTGACAGGCTGGGAGCGCTCCAGAAATGAGGTCAACC 5240
Qy LeuThrHisProValThrLysTyrIleAlaThrCysMetGlnAlaAspLeuGluValMet 1660
Db CTCACCCACCCCAATACCAATACATCATGCGATCGATGTCCGGCTGACCTGGAGGTCTGC 5300
Qy ThrSerThrTrpValLeuAlaGlyValLeuAlaAlaValAlaAlaTyrCysLeuAla 1680
Db ACTAGCACTGGGTGTGGTGGCGAGTCTCTGACGCTCTGCGCGGTATTGCTCCCTGACA 5360
Qy ThrGlyCysValCysIleGlyArgLeuHisIleAsnGlnArgAlaValValAlaPro 1700
Db ACAGCAGTGTGTTCATTTGGGTAGGATTTATTTGTCCGGGAGCGCCGCTATGTTCC 5420
Qy AspLysGluValLeuTyrGluAlaPheAspGluMetGluCysAlaSerArgAlaAla 1720
Db GACAGGAGCTTCTCTACAGAGTTCGATGAAATGGAAGAGTGGCTTCGACCTCCCTCC 5480
Qy LeuIleGluGlyGlnArgIleAlaGluMetLeuLysSerLysIleGlnGlyLeuLeu 1740
Db TACATCGAGCAGGGAATCAGCTCGCCGAGCAATTCAGAGCAAGCGCTCGGGTTACTG 5540
Qy GlnGlnAlaSerLysGlnAlaGlnAspIleGlnProThrValGlnAlaSerTrpProLys 1760
Db CAAACAGCCACCAACAGCGGAGGTGTCTGTCTCCCGTGGTGGAGTCCAAGTGGCGAGCC 5600
Db

1761 ValGluGlnPheThrAlaLysHisMetTrpAsnPhelSerGlyLeuGlnTyrLeuAla 1780
 5601 CTTGAGACATCTTGGGCGAGACACATGCGAATTCATCAGCGGATACAGTACTAGCA 5660
 1781 GlyLeuSerThrLeuProGlyAsnProAlaValAlaSerMetMetAlaPheSerAlaAla 1800
 5661 GGCTTATCCACTCTGCTCGGAACCCCGAATAGCATCATGATGGCATTCACAGCCTCT 5720
 1801 LeuThrSerProLeuSerThrThrThrLeuLeuAsnIleLeuGlyTyrLeu 1820
 5721 ATCAGCAGCGCGCTACACCAAGTACCTCTGTTAAACATCTTGGGGGGGGTGGG 5780
 1821 AlaSerGlnIleAlaProAlaGlyAlaThrGlyPheValValSerGlyLeuValGly 1840
 5781 GCTGCCCACTGCCCGCCCGCCAGCGCGCTTGGCTTCTGGGGCGCGGCATCGCCGCT 5840
 1841 AlaAlaValGlySerIleGlyLeuGlyLysValLeuValAspIleLeuAlaGlyTyrGly 1860
 5841 GCGGCTGTTGGCAGCATAGGCTTGGGAAGGTGCTTGGACATCTCGGCGGTTATGA 5900
 1861 AlaGlyIleSerGlyAlaLeuValAlaPheLysIleMetSerGlyLysProSerMet 1880
 5901 GCAGGAGTGGCGCGCGCTCGTGGCGCTTAAAGTCATGAGCGCGGAGATGCCCTCAC 5960
 1881 GluAspValValAsnLeuLeuProGlyIleLeuSerProGlyAlaLeuValValGlyVal 1900
 5961 GAGGACTGTGTCAATCTTCTCTGTCATCTCTCTCTGCGCCCTGCTGCTGGGCTC 6020
 1901 IleCysAlaAlaIleLeuArgArgHisValGlyProGlyGlyLeuAlaValGlnTrpMet 1920
 6021 GTGTGTGCAGCAATCTGCTGTCACACGCTGGGTCCGGGAGAGGGGCTGTGCGAGTGA 6080
 1921 AsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaProThrHisTyrValThr 1940
 6081 AACCGGCTGATAGCTTGGCTCGCGGGGTAATCATGTTTCCCGCCACGCACTATGCTC 6140
 1941 GluSerAspAlaSerGlnArgValThrGlnLeuLeuGlySerLeuThrIleThrSerLeu 1960
 6141 GAGAGCGACCCCGCGCGGTGTTACTCAGATCTCTCCAGCCCTTACCATCACTAGCTG 6200
 1961 LeuArgArgLeuHisAsnTrpIleThrGluAspCysProIleProCysGlyGlySerTrp 1980
 6201 CTGAAGAGGCTCCACAGTGGATTAATGAAGACTGCTCCACACCGTGTTCGGCTCGTGG 6260
 1981 LeuArgAspValTrpAspTrpValCysThrIleLeuThrAspPheLysAsnTrpLeuThr 2000
 6261 CTAAGGAGTGTGGGACTGATATGACCGGTGTGACTGACTTCAAGACCTCGCTCCAG 6320
 2001 SerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTyrLys 2020
 6321 TCAAGCTCTGCGCGAGCTTACTGAGTCCCTTTTCTGTCGCCAACCGGGTTAAAG 6380
 2021 GlyValTrpAlaGlyThrGlyIleMetThrThrArgCysProCysGlyAlaAsnIleSer 2040
 6381 GGAGTCTGGGCGGAGAGCATCATGCAACACCATCGCCATGTGGGACACAGATCACC 6440
 2041 GlyAsnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTrp 2060
 6441 GGACATGTCAAAACCGTTCATGAGGATCGTGGGCTTAAGACCTTCGACGACACAGCTGG 6500
 2061 GlnGlyThrPheProIleAsnCysTyrThrGluGlyGlnCysValProLysProAlaPro 2080
 6501 CATGGAAATCTCCCATCAACGATACACACCGGCGCCCTGCACACCTCTCCAGCGCCA 6560
 2081 AsnPheLysValAlaIleTrpArgValAlaAlaSerGlyTyrAlaGluValThrGlnHis 2100
 6561 AACTATTCTAGGCGCTGTGGCGGTGGCGCTGAGGAGTACGTGGAGGTTCACCGCGGTG 6620
 2101 GlySerTyrHisTyrIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeu 2120
 6621 GGGGATTTCCACTACGTGACGGGCGATGACCACTGACAACTGAAAGTGCCTATCCAGGTT 6680
 2121 ProSerProGluPhePheSerTrpValAspGlyValGlnIleHisArgPheAlaProThr 2140

6681 CCGGCTCTCGAATCTTCTCGAGGTGACGAGTGGCGTTGCACAGGTACGCTCGCGG 6740
 2141 ProLysProPhePheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal 2160
 6741 TGCAGGCTCTCTTACGGGAGAGGTATCAATTCAGGTTCGGGCTCAACAATACCTGGT 6800
 2161 GlySerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThr 2180
 6801 GGTTCACAGCTACCATGCGAGCCGAAACCGGATGTAGCAGTGTCTCATTCCATGCTACC 6860
 2181 AspProSerHisIleThrAlaGluThrAlaAlaArgArgLeuAlaArgGlySerProPro 2200
 6861 GACCCTCCACATCACAGCAACACGCTAGCGTAGTGTGGCCAGCGGGTCTCCCCC 6920
 2201 SerGluAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThr 2220
 6921 TCTTTGGCAGCTCTTTCAGCTAGCCAGTGTCTGCGCTTCTTGAAGCGGACATGCAC 6980
 2221 ThrHisGlyLysAlaTyrAspValAspMetValAspAlaAsnLeuPhe 2236
 6981 ACCACCATGCTCTCCGACCGCTGACCTCATCGAGGCCAACCTCTGTGGCGGAGGAG 7040
 2237 MetGlyGlyAspValThrArgIleGluSerGlySerLysValValValLeuAspSerLeu 2256
 7041 ATGGCGCGGAACATCACCCGCGTGGAGTCGGAGAACAAAGTGTGTCTCTGACCTTTC 7100
 2257 AspProMetValGluGluArgSerAspLeuGluProSerIleProSerGluTyrMetLeu 2276
 7101 GACCCTCTCGAGCGGAGGAGATGAGAGGAGTATCCGTTCCGCGGAGATCTCGCGG 7160
 2277 ProLysLysArgPheProAlaLeuProAlaTyrAlaArgProAspTrpAsnProPro 2296
 7161 AAATCAAGAAGTTCCTCGCAGCGATGCCATCTGGGGCGGCCCGATTAACAACCTCCA 7220
 2297 LeuValGluSerTrpLysArgProAspTrpGlnProAlaThrValAlaGlyCysAlaLeu 2316
 7221 CTGTAGTAGTCTCGAAGGACCCGACCTAGCTCCCTCGGTGTGTGACGGTCCCGTTG 7280
 2317 ProProProArgLysThrProThrProProProArgArgArgArgThrValGlyLeuSer 2336
 7281 CCACCTATCAAGAGGCTCCCAATACCATCCACCTCCAGGAGAAAGAGAGCGGTTCCT 7340
 2337 GluAspSerIleGlyAspAlaLeuGlnLeuAlaIleLysSerPheGlyGlnProPro 2356
 7341 GAGTCTCTGCTCTTCTGCTTACGGAGCTCGCTACTAAGACCTTCGCGC 7391
 2357 ProSerGlyAspSerGlyLeuSerThrGlyAlaGlyAlaAlaAspSerGlySerGlnThr 2376
 7392 -----AGCTCCGAATCATCGGCGTCCAGCGGCTCGACAGCGGACCGGCGACC 7430
 2377 Pro-----ProAspGluLeuAla-----LeuSerGluThrGlySerIleSer 2390
 7431 GCGCTTCTGACAGCGCTCCGACGAGGTGACAAAGGATCCGAGCTTGAAGTCTGCTAC 7490
 2391 SerMetProProLeuGluGlyLeuGlyAspProAspLeuGluProGluGlnValGlu 2410
 7491 TCAATGCCCCCTTGGAGGGGAAACCGGGGAGCCCCGATCTC----- 7532
 2411 ProGlnProProGlnGlyValAlaAlaProGlySerAspSerGlySerTrpSer 2430
 7533 -----AGTACGGGCTTGGTCT 7550
 2431 ThrCysSerGluGlu-----AspAspSerValValCysCysSerMetSerTrpTrpThr 2449
 7551 ACCGTGAGGAGGAGTGTAGTGGAGTGTCTGCTGCTCAATGTCTTACATGCTACATG 7610
 2450 GlyAlaLeuIleThrProCysSerProGluGluGluLysLeuProIleAsnProLeuSer 2469
 7611 GCGGCTTGTATCAGCCATGCTGCGGAGGAGAAAGACAGCTGCCATCAACCGGTGAGC 7670
 2470 AsnSerLeuLeuArgTyrHisAsnLysValTyrCysThrThrThrLysSerAlaSerLeu 2489

Db 7671 AACTCTTTCTGCGCCACCATTAACATGGTTTATGCGCAACATCTCGCAGCGCAGGCGCTG 7730
Qy 2490 ArgAlaLysLeuValThrPheAspArgMetGlnValLeuAspSerTyrTyrAspSerVal 2509
Db 7731 CGCGAGAAGAGTCACTTTTCAGACAGCTGCAAGTCTCTGAGCAGCACCCTACCGGAGCGT 7790
Qy 2510 LeuLysAspIleLysLeuAlaLysValThrAlaArgLeuLeuThrMetGluGlu 2529
Db 7791 CTCAGAGGAGTAAAGCGAGCGTCCACAGTTAAGGCTAAAGCTCTCTATCCGTAGAGGA 7850
Qy 2530 AlaCysGlnLeuThrProHisSerAlaArgSerLysTyrGlyPheGlyAlaLysGlu 2549
Db 7851 GCGTGAAGCTACGCGCCCAATTCGCGCAATCCAGTTTGGCTATGGCGCAAGGAC 7910
Qy 2550 ValArgSerLeuSerGlyArgAlaValAanHisIleLysSerValTyrLysAspLeuLeu 2569
Db 7911 GTCCGGAACCTATCCAGCAGGCGCTTACCACTCCCTCGTGTGAGAGACTTGCTG 7970
Qy 2570 GluAspSerGluThrProLysLeuThrIleMetAlaLysAsnGluValPheCysVal 2589
Db 7971 GAAGACACTGTGACACCAATTCACACCATGCGCAAAATGAGGTTTCTGTGTC 8030
Qy 2590 AspProThrLysGlyLysLysAlaAlaArgLeuIleValTyrProAspLeuGlyVal 2609
Db 8031 CAACAGAGAAAGGAGCGCTAAGCGCCGCTTATCGTATCCAGATCTGGGAGTC 8090
Qy 2610 ArgValCysGluLysMetAlaLeuTyrAspIleThrGlnLysLeuProGlnAlaValMet 2629
Db 8091 CGTGATGCGAGAGATGCGCTCTATGATGGTCTCCACCTTCTCAGGTCTGTGATG 8150
Qy 2630 GlyAlaSerTyrGlyPheGlnTyrSerProAlaGlnArgValGluPheLeuLysAla 2649
Db 8151 GCGTCTCTATACGATTCAGTACTCTCTGGCGCAGCGAGTCCGAGTCTCGTGAATACC 8210
Qy 2650 TrpAlaGluLysAspProMetGlyPheSerTyrAspThrArgCysPheAspSerThr 2669
Db 8211 TGGAAATCAAGAAAGAAACCCATGGGCTTTTATATGACACCTCGCTGTTCGACTCAACG 8270
Qy 2670 ValThrGluArgAspIleArgThrGluGluSerIleTyrArgAlaCysSerLeuProGlu 2689
Db 8271 GTCCAGCAGACGACATCCGTTGTGAGGAGTCAATTTACCAATGTTGCACTTGGCCCCC 8330
Qy 2690 GluAlaHisThrAlaIleHisSerLeuThrGluArgLeuTyrValGlyGlyProMetPhe 2709
Db 8331 GAAGCCAGACAGCCATAAATCGCTCACAGAGCGCTTTATATCGGGGCTCTCTGACT 8390
Qy 2710 AsnSerLysGlyGlnThrCysGlyTyrArgArgCysArgAlaSerGlyValLeuThrThr 2729
Db 8391 AATTCAAAAGGGCAGAACTGCGGTTATTCGCGGTGCGCGCGCGCGCTGCTGACGACT 8450
Qy 2730 SerMetGlyAsnThrIleThrCysTyrValLysAlaLeuAlaCysLysAlaAlaGly 2749
Db 8451 AGCTGGGTAAACCTCATGTTACTTGAAGGCTCTGAGCGCTGTGAGCTGGGAAG 8510
Qy 2750 IleIleAlaProThrMetLeuValCysGlyAspAspLeuValIleSerGluSerGln 2769
Db 8511 CTCCAGGACTGACGAGTCTCGTGAACGAGAGCAGCTCTGCTATCTGTGAAGCGCG 8570
Qy 2770 GlyThrGluLysAspGluArgAsnLeuArgAlaPheThrGluAlaMetThrArgTyrSer 2789
Db 8571 GGAACCCAAAGAGCAGCGCGGAGCTTACGAGTCTTCCAGCGAGCTATGACTAGTACTCC 8630
Qy 2790 AlaProGlyAspProArgProGluTyrAspLeuGluLeuIleThrSerCysSer 2809
Db 8631 GCGCCCCCGGAGCCCGCCCAACCAAGATAGACTTGAGCTGATACATCATGTTCC 8690
Qy 2810 SerAsnValSerValAlaLeuGlyProGlnGlyArgArgTyrTyrLeuThrArgAsp 2829
Db 8691 TCCAATGTGCTGCGCCCGCAGATGATCAGGCAAAAGGGTGTACTACCTCACCGCTGAT 8750
Qy 2830 ProThrThrProIleAlaArgAlaAlaTrpGluThrValArgHisSerProValAsnSer 2849
Db 8751 CCCACCACCCCTAGCAGCGGCTGCGTGGAGACAGCTAGACACACTCCAGTTAACTCC 8810

Qy 2850 TrpLeuGlyAsnIleIleGlnTyrAlaProThrIleTrpAlaArgMetValLeuMetThr 2869
Db 8811 TGGCTAGGCAACATTATTATGATGATCGCCACTTTGTGGCAAGGATGATTCTGATGACT 8870
Qy 2870 HisPhePheSerIleLeuMetAlaGlnAspThrLeuAspGlnAsnLeuAsnPheGluMet 2889
Db 8871 CACTTCTTCTCCATCTTCTAGCGGAGGCAACTTGAAGAGCCCTCGACTGCCAGATC 8930
Qy 2890 TyrGlyAlaValTyrSerValSerProLeuAspLeuProAlaIleIleGlnArgLeuHis 2909
Db 8931 TAGCGGCGCTGTACTTCCATGAGCATTGACCTACTCTCAGTCACTTTGAAGCACTCCAT 8990
Qy 2910 GlyLeuAspAlaPheSerLeuHisThrTyrThrProHisGluLeuLeuThrArgValAlaSer 2929
Db 8991 GGCCTTAGCGCATTTTCACTCCATGTTACTCTCCAGGTGAGATCAATAGGGTGGCTTCA 9050
Qy 2930 AlaLeuArgLysLeuGlyAlaProProLeuArgAlaTrpLysSerArgAlaArgAlaVal 2949
Db 9051 TGCCTCAGGAACTTTGGGGTACCACTTGGAGTCTGGAGACATCGGCGCAGGAGCGTC 9110
Qy 2950 ArgAlaSerLeuIleSerArgGlyGlyArgAlaAlaValCysGlyArgTyrLeuPheAsn 2969
Db 9111 CCGCTAGGCTACTGCTCCAGGAGGAGGCGGCGCCTTGTGGCAATACCTCTCTCAAC 9170
Qy 2970 TrpAlaValLysThrLysLeuLysLeuThrProLeuProGluAlaArgLeuLeuAspLeu 2989
Db 9171 TGGCGAGTAAACCACTTAACTCACTCCAACTCCCGGCTGCGTCCCGGCTGGACTTG 9230
Qy 2990 SerSerTrpPheThrValGlyAlaGlyGlyAspIleTyrHisSerValSerArgAla 3009
Db 9231 TCCGCGTGGTGTGTTGCTGTGTACAGCGGGGAGACATATATCACAGCCTGTCTCGTGCC 9290
Qy 3010 ArgProArgLeuLeuPheGlyLeuLeuLeuPheValGlyValGlyLeuPheLeu 3029
Db 9291 CGACCCCGTGGTTCATGCTGCTGCTACTCTCTACTTCTGTAGGGGTAGGCACTACCTG 9350
Qy 3030 LeuProAlaArg 3033
Db 9351 CTCCCAACCGA 9362

RESULT 10

US-08-384-616-1
; Sequence 1, Application US/08384616
; Patent No. 5847101
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Armstrong, Westerman, Hattori, Mclelland &
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/384,616
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996

FILING DATE: 02-OCT-1991
 APPLICATION NUMBER: JP 2-167466
 FILING DATE: 25-JUN-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 2-230921
 FILING DATE: 31-AUG-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 2-305605
 FILING DATE: 09-NOV-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/635,451
 FILING DATE: 28-DEC-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Stevens-Smith, Theresa M.
 REGISTRATION NUMBER: 36,281
 REFERENCE/DOCKET NUMBER: 900703B
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 659-2930
 TELEFAX: (202) 887-0357
 TELEX: 440142
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 9416 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA from genomic RNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 333..9362
 US-08-384-616-1

Alignment Scores:
 Pred. No.: 0
 Score: 11984.50
 Length: 9416
 Matches: 2175
 Percent Similarity: 83.02%
 Conservative: 352
 Best Local Similarity: 71.45%
 Mismatches: 472
 Query Match: 74.35%
 Indels: 45
 Gaps: 8

US-09-980-559-2 (1-3033) x US-08-384-616-1 (1-9416)

1 MetSerThrAsnProGlnArgLysThrLysArgAsnThrAsnArgProGln 20
 333 ATGAGCAGCAATCTTAACCTCAAGAAACCAACAGTACCAACCGCCGCCACAG 392
 21 AspValLysPheProGlyGlyGlnIleValGlyValLysLeuLeuProArg 40
 393 GACGTCAAGTTCCCGGGCGGTGTCAGATCGTTGGTGGAGTTTACCTGTTCGCGCAGG 452
 41 GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly 60
 453 GGCCTCCAGGTGGGTGTCGCGGGCCCGCAGGAAGACTTCGAGCGGTTCGCAACCTCGTGA 512
 61 ArgArgGlnProLysPheAspArgSerThrGlyLysSerThrGlyLysProGly 80
 513 AGCGGACACCTATCCCAAGGCTCGCCGCCCGCAGGACCTCGGGCTAGCCCGGG 572
 81 TyrProTrpProLeuTyrGlyAsnGluGlyLeuGlyTrpAlaGlyTrpLeuLeuSerPro 100
 573 TACCTTGGCTCTCTATGCAATGAGGCTTAGGGTGGGCGAGGATGGCTCTGTCAACC 632
 101 ArgGlySerArgProSerTrpGlyProAsnAspProArgHisArgSerArgAsnValGly 120
 633 CGCGGCTCCCGGCTAGTTGGGGCCCGCAGGACCCCGCGGTAGGTGCGGTAAATTTGGGT 692
 121 LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTrpIleProValVal 140
 693 AAGGTATCATACCTCTACATCGGCTTCGCGAICTCAAGGGGTACATTCGCTCGTC 752
 141 GlyAlaProLeuGlyGlyValAlaArgAlaLeuAlaHisGlyValArgValLeuGluAsp 160
 753 GCGGCCCTTGGGGCGGTGTCAGGCGCCCTGGCACATGCTGTCCGGGTCTCTGGAGGAC 812

161 GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla 180
 813 GCGTGAACACTATGCAACAGGAACTGCGCGGTGCTCTTTCTTCTATCTCTCTTGGCT 872
 181 LeuLeuSerCysIleThrThrProValSerAlaAlaGluValLysAsnIleSerThrGly 200
 873 CTGCTGTCTGCTGACCAACCCAGCTTCCGCTTACGAAGTGCACAAAGTGTCCGGGATA 932
 201 TyrMetValThrAsnAspCysThrAsnAspSerIleThrTrpGlnLeuGlnAlaVal 220
 933 TATCATGTCAACAGCACTGCTCCCAACGAACATGTTGATGAGGCGAGCGGACTTGATC 992
 221 LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTrpIle 240
 993 ATGCACTACTCTGGGTGCTGCTGCTGCGGAGGCACTCTCTCCGCTGCTGGGTA 1052
 241 ProValSerProAsnValAlaValGlnArgProGlyAlaLeuThrGlnGlyLeuArgThr 260
 1053 GCGCTCACTCCCAACGCTCGAGCAGAAAGTCCACATCCCAACGACGATACGACGC 1112
 261 HisIleAspMetValValMetSerAlaThrLeuCysSerAlaLeuTyrValGlyAspLeu 280
 1113 CACGTGATCTGCTGCTGCGGCGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1172
 281 CysGlyGlyValMetLeuAlaAlaGlnMetPheIleValSerProGlnHisHisTrpPhe 300
 1173 TCGGATCTGTTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1232
 301 ValGlnAspCysAsnCysSerIleTyrProGlyThrIleThrGlyHisArgMetAlaTrp 320
 1233 TTACAGGACTGTAACTGCTCAATTTATCCCGGCACTGTGCTGGGTCACTGCTGCTGCTGCT 1292
 321 AspMetMetMetAsnTrpSerProThrAlaThrMetIleLeuAlaTyrAlaMetArgVal 340
 1293 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1352
 341 ProGluValIleIleAspIleIleSerGlyAlaHisTrpGlyValMetPheGlyLeuAla 360
 1353 CCACAAAGCGCTGCTGACATGCTGGGGGGGCGGAGTCTGGGGAGTCTGGGGGCTTGGC 1412
 361 TyrPheSerMetGlnGlyAlaTrpAlaLysValValValIleLeuLeuLeuAlaGly 380
 1413 TACTATTCCATGCGGGGAACTGGGCTAAGTTCTGATGATGATGATGATGATGATGATGATG 1472
 381 ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr 400
 1473 GTTGACGGGATACCCACGCTGACGGGGGGGCGGAGCCAAACCAACCAACCAACCAACCA 1532
 401 SerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTrp 420
 1533 TCATGTTGCGAAGTGGGCGCTCTGAAATAATCCAGCTTATAACCAACCAACCAACCAACCA 1592
 421 HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer 440
 1593 CACATCAACAGGACTGCCCTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 1652
 441 LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg 460
 1653 CTGTTCTACACATAGTTTCAACTGCTCGGGGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 1712
 461 SerIleGluAlaPheArgValGlyTrpGlyAlaLeuGlnTyrGluAspAsnValThrAsn 480
 1713 ACCATTGACAGTTCCAGCAGGATGGGTCCCACTTACTTATGCTAG-----TCTAGC 1766
 481 ProGluAspMetArgProTyrCysTrpHisTyrProProArgGlnCysGlyValValSer 500
 1767 AGATCAGACAGAGGCGCATTTGCTGGCACTACCCACCTCCCAACATGATCATCTACCT 1826
 501 AlaLysThrValCysGlyProValTyrCysPheThrProSerProValValGlyThr 520
 1827 GCGTGGAGGTGCGGGCGGCTGCTGCTTACCCCAAGCCCTGCTGCTGCTGCTGCTGCTGCTG 1886

```
QY 521 ThrAspArgLeuGlyAlaProThrTyrThrTriphGlyAlaAsnGluThrAspValPheLeu 540
Db 1887 ACCGATCGTTTCGGTGTCCCTACGATAGATGGGGGAGAACGAGACTGACGCTGCTG 1946
QY 541 LeuAsnSerThrArgProProLeuGlySerTriphPheGlyCysThrTriphMetAnSerSer 560
Db 1947 CTCAACACACGGCGCGCGCGCGCAAGCAACTGGTTCGGCTGCACATGGATGAATGACACC 2006
QY 561 GlyTyrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAlaSer 580
Db 2007 GGGTTACCAAGACATGTGGGGGCCCCCGGTAAACATCGGGGGGTGCGCAAC----- 2060
QY 581 ThrAspLeuLeuCysProThrAspCysPheArgLysHisProAspThrThrLysLeuLys 600
Db 2061 AACACCCCTGACCTGCCCCACGGACTGCTCCGGAAGCACCCCGAGGCTACCTACACAAA 2120
QY 601 CysGlySerGlyProTriphLeuThrProArgCysLeuLeuLeuAspTyrProTyrArgLeuTriph 620
Db 2121 TGTGGTTTCGGGGGCTTGGCTGCACACCTAGGTGATGGTTGACTATCCATACAGGCTCTGG 2180
QY 621 HisTyrProCysThrValAsnTyrThrIlePheLysIleArgMetTyrValGlyGlyVal 640
Db 2181 CATTACCCCTGCACTGTATACCTTTTACCATCTTCAAGSTTAGGATGATGTGGGGGGGTG 2240
QY 641 GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp 660
Db 2241 GAGCACAGGCTCAATGCTGATCAATGACCAATGGACCCGAGGAGCGTTGTGACTTGGAGGAC 2300
QY 661 ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGluTriphAlaIleLeuPro 680
Db 2301 AGGGATAGGCGGAGGCTCAGCCGCTGCTGTCTACAAACAGAGTGGCAGGTACTGCC 2360
QY 681 CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisLeuLeuHisGlnAsnIle 700
Db 2361 TGTTCTTACCACTTACCAGCTCTGTCCACTGGCTTGATTCACCTCCATCAGAACATC 2420
QY 701 ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrLysTyrIleValArgTriph 720
Db 2421 GTGGAGCTGCAATACCTATACGTTATAGGTGACGGTGTCTCTCTTTCGCAATCAATGG 2480
QY 721 GluTriphValIleLeuLeuPheLeuLeuAlaAspAlaArgValCysAlaCysLeuTriph 740
Db 2481 GAGTATGTCTGTGTTCTTCTCTCTAGCGGACGACGTCTGTCTGTCTGTCTGTCTGTCTGT 2540
QY 741 MetLeuLeuLeuGlyGlnAlaGluAlaAlaLeuGluLysLeuValIleLeuHisAla 760
Db 2541 ATGATGCTGTGATAGCCAGCGCGGCGCGCTTGGAGAACCTTGGTGGTCTCAATTCG 2600
QY 761 AlaSerAlaAlaSerCysAsnGlyPheLeuTyrPheValIlePheValAlaAlaTriph 780
Db 2601 GCGTCTGTGCGCGGCACATGCGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2660
QY 781 TyrIleLysGlyArgValProLeuAlaThrTyrSerLeuThrGlyLeuTriphSerPhe 800
Db 2661 TACATCAAGAGCGAGGCTGGTCCCTGGGGGCGACATATGCTCTTTATGGCGTGTGGCGCGCTG 2720
QY 801 SerLeuLeuLeuAlaLeuProGlnAlaTyrAlaTyrAspAlaSerValHisGly 820
Db 2721 CTCCTGTCTGTGTGGCATTACACCGCGAGCTTACGCCATGACCGGGAGATGGGTGCA 2780
QY 821 GlnIleGlyAlaAlaLeuValMetIleThrLeuPheThrLeuThrProGlyTyrLys 840
Db 2781 TCGTGGGAGCGGGTTTTTGGGTCTGGTACTCTGTACTTGTGTGTGTGTGTGTGTGTGTGTGT 2840
QY 841 ThrLeuLeuSerArgPheLeuTriphTriphLeuCysTyrLeuLeuThrLeuGlyGluAlaMet 860
Db 2841 GTGTTCCTCGTAGGCTCATATGGTGTGTACAAATATTTTACCACAGAGCGGAGCGGAC 2900
QY 861 ValGlnGluTriphAlaProProMetGlnValArgGlyGlyArgAspGlyIleIleTriphAla 880
Db 2901 TTACATGTGTGGATCCCCCCCCCTCAACGCTCGGGGAGCGCGCATGCCATCTCTCTCTCT 2960
QY 881 ValAlaIlePheTyrProGlyValValPheAspIleThrLysTriphLeuLeuAlaValLeu 900
```


4041 AGTAAAGTCCGCGCTGCATATGACAGCCAAAGGTGCTCGTCTCAATCCG 4100
 1261 SerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerIleAlaHisGlyIleAsnPro 1280
 4101 TCGGTTGCCGCTACCTTAGGGTTTGGGGGTATATGCTTAAGGACACACGGTATTTGACCCC 4160
 1281 AsnIleArgThrGlyValArgThrValThrThrGlyAlaProIleThrTyrSerThrTyr 1300
 4161 AACATCAGAATCGGGTAGGACCAATACACAGGCGCCCGCTCACATACCTACTAT 4220
 1301 GlyLysPheLeuAlaAlaGlyGlyCysAlaGlyAlaTyrAspIleIleIleCysAsp 1320
 4221 GGCAGGTTCTCCGATGCTGCTGCTGCGGCGGCTTATGACATCATATAATATGAT 4280
 1321 GluCysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGlnAla 1340
 4281 GAGTGCATTTCAACTGACTCGACTACAACTCTGGGCATCGGCAGCTCTGGACCAAGCG 4340
 1341 GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProGlySerValThr 1360
 4341 GAGACGGCTGGAGCGGCTTGTGCTGCTCGCCACCGCTACGCTCGGGATCGGTCAAC 4400
 1361 ThrProHisProAsnIleGluGluValAlaLeuGlyGlnGluGlyGluIleProPheTyr 1380
 4401 GTGCCACACCAACATCGAGGAGTGGCCCTGCTTAATACTCGAGAGATCCCGCTTCTAT 4460
 1381 GlyArgAlaIleProLeuSerTyrIleLysGlyGlyArgHisIleLeuIlePheCysHisSer 1400
 4461 GGCAGGACCATCCCAATTAAGCCATCAGGGGGGAGGCAATCTCAATTTCTCTCATTC 4520
 1401 LysLysLysCysAspGluLeuAlaAlaAlaLeuArgGlyMetGlyLeuAsnSerValAla 1420
 4521 AAGAGAGTGGCAGGAGCTCGCGCAAGCTGTGAGGCTCGGATCAAGCTGTGGCG 4580
 1421 TyrTyrArgGlyLeuAspValSerValIleProThrGlnGlyAspValValValAla 1440
 4581 TATTACCGGGGCTCGATGTGCTGCTCATACCAACTATCGAGAGCTGCTGTGCTGGCA 4640
 1441 ThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnVal 1460
 4641 ACAGACGCTCTGATACGGGCTTATGAGGGGAGCTTGTGACTGAGTGTGACTGTAAACA 4700
 1461 AlaValThrGlnValValAspPheSerLeuAspProThrPheThrIleThrGlnIle 1480
 4701 TGTGTACCCAGACAGCTGACTTGTGATCCGCTTACCATTTGAGACGACGACC 4760
 1481 ValProGlnAspAlaValSerArgSerGlnArgArgGlyArgThrGlyArgGlyArgLeu 1500
 4761 GTGCTCAAGACGCGTGTGCTGCTGCGCTCGAGCGGGGTAGGACTGCGAGGGGTAGGAGA 4820
 1501 GlyIleTyrArgTyrValSerThrGlyGluArgAlaSerGlyMetPheAspSerValVal 1520
 4821 GGCATCTACAGGTTTGTGACTCCGGGAGAACGGCCCTCGGGCAGTTTGTGATCTCTGCTG 4880
 1521 LeuCysGluCysTyrAspAlaGlyAlaAlaTyrTyrGluLeuThrProSerGluThrThr 1540
 4881 CTGTGTGATGCTATGACGGGGCTGTGCTGTGATGAGCTCACCCCGGCGGAGACTCG 4940
 1541 ValArgLeuArgAlaTyrPheAsnThrProGlyLeuProValCysGlnAspHisLeuGlu 1560
 4941 GTTAGTTGCGGGCTACCTGACACACACAGGAGTGGCCGCTTGGCCAGGACCACTGGAG 5000
 1561 PheTyrGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThr 1580
 5001 TTCTGGGAGAGTGTCTTTCACAGGCTCACCACATATAGATGACACTTCTTGTCCAGACC 5060
 1581 LysGlnSerGlyGluAsnPheAlaTyrLeuThrAlaTyrGlnAlaThrValCysAlaArg 1600
 5061 AAGCAGCAGGAGGACAACTTCCCTACTCTGGTAGCATACCAAGCCAGGTGTGGCCAGG 5120
 1601 AlaLysAlaProProProSerTyrAspValMetTyrLysCysLeuThrArgLeuLysPro 1620
 5121 GCTCAGGCCCCACCTCATCGGATCAAAATGTGGAAGTGTCTCATACGCTGAAACCT 5180

1621 ThrLeuValGlyProThrProLeuLeuTyrArgLeuGlySerValThrAsnGluValThr 1640
 5181 ACAGTGCAGGGCCAAACCCCTTGTGTACAGGCTGGAGCGCTCCAGAAATGAGGTCAACC 5240
 1641 LeuThrHisProValThrLysTyrIleAlaThrCysMetGlnAlaAspLeuGluValMet 1660
 5241 CTCACCCACCCCAATACCAAAATACATCATGTCATGTCGCTGACCTGAGGTGCTC 5300
 1661 ThrSerThrTyrValLeuAlaGlyGlyValLeuAlaAlaValAlaTyrCysLeuAla 1680
 5301 ACTAGACCTTGGGTGTGTGGCGGAGTCTTGTGACGCTCTGCGCGGTATTTGCCGTGACA 5360
 1681 ThrGlyCysValCysIleIleGlyArgLeuHisIleAsnGlnArgAlaValValAlaPro 1700
 5361 ACAGGAGTGTGCTCATTTGTGGTAGGATTATCTTGTCCGGAGGCGCGCCATTGTTCC 5420
 1701 AspLysGluValLeuTyrGluAlaPheAspGluMetGluGluCysAlaSerArgAlaAla 1720
 5421 GACAGGAGCTTCTCTACCGAGGTTTGTGATGAATGAAGAGTGGCGCTCGACCTCCCT 5480
 1721 LeuIleGluGluGlyGlnArgIleAlaGluMetLeuLysSerLysIleGlnGlyLeuLeu 1740
 5481 TATATCGAGCAGGGAATGAGCTCGCGGAGCAATTCACAGCAAGAGCGCTCGGTTACTG 5540
 1741 GlnGlnAlaSerLysGlnAlaGlnAspIleGlnProThrValGlnAlaSerTyrProLys 1760
 5541 CAACAGCACCAACACAGGAGGCTGCTGCTCCGCTGCTGAGTCCAAGTGGCGAGCC 5600
 1761 ValGluGlnPheTyrAlaLysHisMetTyrAsnPheIleSerGlyIleGlnTyrLeuAla 1780
 5601 CTTGAGACATCTCGGCGAAGCACATGTGAATTTTCATCAGCGGAGTACAGTACTTAGCA 5660
 1781 GlyLeuSerThrLeuProGlyAsnProAlaValAlaSerMetMetAlaPheSerAlaAla 1800
 5661 GCTTATTCACCTGCTGCTGGGAACCCCGCAATGAGCATCATGATGGCATTTACAGCTCT 5720
 1801 LeuThrSerProLeuSerThrThrIleLeuLeuAsnIleLeuGlyGlyTyrLeu 1820
 5721 ATCCAGGCGCGCTCACACCCCAAGTACCTCTCTGTTTAAACATCTTGGGGGTGGTG 5780
 1821 AlaSerGlnIleAlaProProAlaGlyAlaThrGlyPheValValSerGlyLeuValGly 1840
 5781 GCTGCCAACTCGCCCGCCCGCGCTTGGCTTCTGCGGCGCGGCTCGGCTCGGCT 5840
 1841 AlaAlaValGlySerIleGlyLeuGlyLysValLeuValAspIleLeuAlaGlyTyrGly 1860
 5841 GCGGCTGTGTGAGCAGCATAGGCTTGGGAAGGTGCTTGTGACATCTTGGCGGTTATGGA 5900
 1861 AlaGlyIleSerGlyAlaLeuValAlaPheLysIleMetSerGlyGlyLysProSerMet 1880
 5901 GCAGGAGTGGCGCGCTCGTGGCTTTTAAAGTTCATGAGCGGCGAGATGCCCTCCACC 5960
 1881 GluAspValValAsnLeuLeuProGlyIleLeuSerProGlyAlaLeuValGlyVal 1900
 5961 GAGGACCTGCTCAATCTACTTCTGTCATCTCTCTGCGGCGCTGCTGCTGCGGCTC 6020
 1901 IleCysAlaAlaIleLeuArgHisValGlyProGlyGluGlyAlaValGlnTyrMet 1920
 6021 GTGTGTGAGCAATATCTGCTGACACGCTGGGTTCGGGAGAGGGGCTGTGAGTGTGATG 6080
 1921 AsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaProThrHisTyrValThr 1940
 6081 AACCGCTGATAGGTTTCCCTCGCGGGTAAATCATGTTTCCCGCCAGCATATGTGCT 6140
 1941 GluSerAspAlaSerGlnArgValThrGlnLeuLeuGlySerLeuThrIleThrSerLeu 1960
 6141 GAGAGCGACCGCGCGGCTGTGTTACTCAGATCTCTCCAGCTTACCATCACTCACTGCT 6200
 1961 LeuArgArgLeuHisAsnTyrIleThrGluAspCysProIleProCysGlyGlySerTyr 1980
 6201 CTGAAAGAGCTCCACAGTGGATTAAAGAGTGTCTCCACACCGTGTCCGCTCTGTTG 6260

[illegible]

Db 8331 GAAGCCAGACAGCCATAAAATCGCTCACAGACGGCTTTATATCGGGGTCTCTGACT 8390
 Qy 2710 AnSerlysglyGlnThrCysGlyTyrArgAAserGlyValLeuThrThr 2729
 Db 8391 AATTCAAAGGGCAGAACTTCGGTTCATCGCGGTGCGCGAGCGGTGCTGACACT 8450
 Qy 2730 SerMetGlyAsnThrIleThrCysTyrValLysAlaLeuAlaCysLysAlaGly 2749
 Db 8451 AGCTGGGTAAACCTCATCTTACTTGAAGCCTCTGCAGCCTGTCAGCTGGAAG 8510
 Qy 2750 IleAlaProThrMetLeuValCysGlyAspLeuValLysSerGluSerGln 2769
 Db 8511 CTCAGGACTGACGATGCTGCTGAACGAGACGCTCTGCTTATCTGTGAAGCGG 8570
 Qy 2770 GlyThrGluAspGluArgAsnLeuAlaAlaPheThrGluAlaMetThrArgTyrSer 2789
 Db 8571 GGAACCCAGAGACGGCGGACGCTACGAGTCTTCAGGAGGCTATGACTAGTACTCC 8630
 Qy 2790 AlaProGlyAspProArgProGluTyrAspLeuGluLeuIleThrSerCysSer 2809
 Db 8631 GCCCCCCGGGACCGCCGCCACACGATACGACTTGGAGCTGATACATCATCTTCC 8690
 Qy 2810 SerAsnValSerValAlaLeuGlyProGlnGlyArgArgTyrTyrLeuThrArgAsp 2829
 Db 8691 TCCAAATGTCGCTCGCCACGATGCTATCAGGCAAAAGGTGTACTACCTCACCCTGAT 8750
 Qy 2830 ProThrThrProIleAlaAlaAlaTyrGluThrValArgHisSerProValAsnSer 2849
 Db 8751 CCCACACCCCTTAGCAGCGGTGCGTGGAGACAGCTAGACACACTCCAGTTAACTCC 8810
 Qy 2850 TrpLeuGlyAsnIleGlnTyrAlaProThrIleTyrAlaAlaArgMetValLeuMetThr 2869
 Db 8811 TGCTAGGCAACATTATATGATGCGCCACCTTGTGGCAAGGATGATCTTGATGACT 8870
 Qy 2870 HisPhePheSerIleLeuMetAlaGlnAspThrLeuAspGlnAsnLeuAsnPheGluMet 2889
 Db 8871 CACTTCTCTCCATCTCTTAGCGGACGACCACTTGAAAAAGCCCTGGAGCTCCAGATC 8930
 Qy 2890 TyrGlyAlaValTyrSerValSerProLeuAspLeuProAlaIleIleGluArgLeuHis 2909
 Db 8931 TAGGGGCGCTGTACTCCATGAGCCACTTGACCTACCTCAGATCATTTGAACGACTCCAT 8990
 Qy 2910 GlyLeuAspAlaPheSerLeuHisThrTyrThrProHisGluLeuThrArgValAlaSer 2929
 Db 8991 GGCCTTAGCCATTTTCACTCCATGATCTCTCCAGTGAGATCAATAGGTGGCTTCA 9050
 Qy 2930 AlaLeuArgLysLeuGlyAlaProProLeuArgAlaTyrLysSerArgAlaAlaVal 2949
 Db 9051 TGCCTCAGGAACTTGGCGTACCACCTTGGAGTCTGGAGACATCGGGCCAGGAGCTC 9110
 Qy 2950 ArgAlaSerLeuIleSerArgGlyGlyArgAlaAlaValCysGlyArgTyrLeuPheAsn 2969
 Db 9111 CGCGCTAGGTACTGTCTCCAGGAGGAGGCGGCCACTTGTGGCAATACCTCTTCAAC 9170
 Qy 2970 TrpAlaValLysThrLysLeuLysLeuThrProLeuProGluAlaArgLeuLeuAspLeu 2989
 Db 9171 TGGCAGTAAACCAACCTTAACCTCACTCCATCCCGCTCGGCTCGGCTGACTTG 9230
 Qy 2990 SerSerTrpPheThrValGlyAlaGlyGlyAspIleTyrHisSerValSerArgAla 3009
 Db 9231 TCCGGCTGTTGTTGCTGTTTACAGCGGGGAGACATATATCACAGCTGTCTCGTGC 9290
 Qy 3010 ArgProGluLeuLeuPheGlyLeuLeuLeuPheValGlyValGlyLeuPheLeu 3029
 Db 9291 CGACCCCGTGGTTTCATGCTGTGCTACTCTTCTGTAGGGGTAGGATCTACCTG 9350
 Qy 3030 LeuProAlaArg 3033
 Db 9351 CTCCCCAACGA 9362

RESULT 11
 US-08-904-686A-1
 ; Sequence 1, Application US/08904686A

; Patent No. 5998130
 ; GENERAL INFORMATION:
 ; APPLICANT: OKAYAMA, Hiroto
 ; APPLICANT: FUKU, Isao
 ; APPLICANT: MORI, Chisato
 ; APPLICANT: TAKAMIZAWA, Akahisa
 ; APPLICANT: YOSHIDA, Iwao
 ; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
 ; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
 ; NUMBER OF SEQUENCES: 50
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSES: Armstrong, Westerman, Hattori, McIeland &
 ; ADDRESSES: Naughton
 ; STREET: 1725 K St. N.W. Suite 1000
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20006
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.5 in, 1.44mb
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/904,686A
 ; FILING DATE: 01-AUG-1997
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/324,977
 ; FILING DATE: 18-OCT-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 2-167466
 ; FILING DATE: 25-JUN-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 2-230921
 ; FILING DATE: 31-AUG-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 2-305605
 ; FILING DATE: 09-NOV-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/099,706
 ; FILING DATE: 30-JUL-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/769,996
 ; FILING DATE: 02-OCT-1991
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/635,451
 ; FILING DATE: 28-DEC-1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McIeland, Le-Nhung
 ; REGISTRATION NUMBER: 31,541
 ; REFERENCE/DOCKET NUMBER: 900703G
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 659-2930
 ; TELEFAX: (202) 887-0357
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 9416 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA from genomic RNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 333..9362
 ; US-08-904-686A-1
 ;
 ; Alignment Scores:
 ; Pred. No.: 0 Length: 9416
 ; Score: 11984.50 Matches: 2175
 ; Percent Similarity: 83.02% Conservatives: 352
 ; Best Local Similarity: 71.45% Mismatches: 472
 ; Query Match: 74.35% Indels: 45
 ; DB: 2 Gaps: 8

1	MetSerThrAnProLysProGlnAArgLysThrLysArgAsnThrAsnArgArgProGln	20
333	ATGAGCAGCAATCCTAAACCTCAAGAAACCAACAAAGTAAACCAACCGCGCGCCACAG	392
21	AspValLysPheProGlyGlyGlyGlnIleValGlyGlyValTyLysLeuLeuProArgArg	40
393	GACGTCAAGTTCCCGGCGGTGTAGATCGTTGTGGAGTTTACCTGTTCGCGGCGAGG	452
41	GlyProArgLeuGlyValArgAlaThrAArgLysThrSerGluArgSerGlnProArgGly	60
453	GGCCCCAGGTGGGTGTGCGCGCGCCAGGAGACTTCCGAGCGGTGCGCAACCTCGTGG	512
61	ArgArgGlnProLysAspArgArgSerThrGlyLysSerTrpGlyLysProGly	80
513	AGGCGACAACTATCCCAAGGCTCGCGCGCCGAGGAGGAGCATTGGGCTCAGCGCCGG	572
81	TyrProTrpProLeuTyGlyAsnGluGlyLeuGlyTrpAlaGlyTrpLeuLeuSerPro	100
573	TACCTTTGGCTCTCTATGCAATGAGGCTTAGGGTGGGCAAGTAGGCTCTCTGTCAACC	632
101	ArgGlySerArgProSerTrpGlyProAsnAspProArgHisArgSerArgAsnValGly	120
633	CGCGGCTCCGGCTAGTTGGGCGCCCAACGAGCCCGCGCGTAGGTCGGTAATTTGGT	692
121	LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTrpIleProValVal	140
693	AAGGTATCATGATACCTTCACATCGGCTTCGCGCATCTCATGGGTACATTCGCTCGT	752
141	GlyAlaProLeuGlyValAlaAArgAlaLeuAlaHisGlyValArgValLeuGluAsp	160
753	GGCGCCCCCTGGGGGCGCTGCAGGGCCCTGGACATGGTGTCCGGGTTCTGGAGGAC	812
161	GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla	180
813	GGCGTGAACATATGCAACAGGGAATCTGCCCGGTGCTCTTTTCTATCTCTCTGGCT	872
181	LeuLeuSerCysIleThrThrProValSerAlaAlaGluValLysAsnIleSerThrGly	200
873	CTGCTGTCTCTGCTGACCCACCCAGCTTCGCTTACGAAGTGCACACGCTCGCGGATA	932
201	TyrMetValThrAsnAspCysThrAsnAspSerIleThrTrpGlnLeuGlnAlaVal	220
933	TATCATGTCAGAAAGACTGCTCCAACGCAAGCATTTGTATAGGCGAGCGGACTTGATC	992
221	LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTrpIle	240
993	ATGCATACTCTGGGTGCGTCCCTCGGTTCCGGAAGCAACTCTCTCCCGTGTCTGGGTA	1052
241	ProValSerProAsnValAlaValGlnAArgProGlyAlaLeuThrGlnGlyLeuArgThr	260
1053	CGCGTCACTCCCAAGCTCGCAGCAGCAAGAAAGTCACCATCCCCACACAGCATACGACG	1112
261	HisIleAspMetValValMetSerAlaThrLeuCysSerAlaLeuTyValGlyAspLeu	280
1113	CACGTGATCTGTCTGTTGGGGCGGTCTTCTTCTTGTTCGCTATGTAGTGTGGGGGACCTC	1172
281	CysGlyGlyValMetLeuAlaAlaGlnMetPheIleValSerProGlnHisIleTrpPhe	300
1173	TGCGGATCTGTTTTCTCGTCTCTCAGCTGTTTCCCTTCTCGCCTCGCGCATGTGACA	1232
301	ValGlnAspCysAsnCysSerIleTyProGlyThrIleThrGlyHisArgMetAlaTrp	320
1233	TTACAGACTGTAACTGCTCAATTTATCCCGGCCAATGTGTGGGTCACCGTATGGCTTGG	1292
321	AspMetMetMetAsnTrpSerProThrAlaThrMetIleLeuAlaTyAlaMetArgVal	340
1293	GACATGATGTAACCTGGTCGCCCAACACAGCCCTAGTGGGTGCGCAGTTACTCCGATC	1352
341	ProGluValIleLeuAspIleLeuSerGlyAlaHisIleTrpGlyValMetPheGlyLeuAla	360

1353	CCACAAGCGGTCTGGACATGGTGGCGGGGGCCCACTGGGGAGTCTCTGGCGGGCCTTGCC	1411
361	TyrPheSerMetGlnGlyAlaTrpAlaLysValValValIleLeuLeuLeuAlaGly	380
1413	TACTATTCCATGGCGGGGAACTGGGCTTAAGGTTCTGATTGTGACTACTATTTTCTGCG	1472
381	ValAspAlaAArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr	400
1473	GTTGACGGGGATATACCACCGTGACAGGGGGGGCGCAAGCCAAACACCAACAGGCTCGTG	1532
401	SerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTrp	420
1533	TCCATGTTCCGAAGTGGGCGGCTCTCAGAAATCCAGCTTATAAACACCAATGGGAGTTGG	1592
421	HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer	440
1593	CACATCAACAGGACTGCCCTGAACGTGAATGACTCTCTCCAGACTGGGTTTCTTCCGCGG	1652
441	LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg	460
1653	CTGTTCTACACACATAGTTTCAACTCGTCCGGGTGCCAGAGCGCATGGCCAGTCCGCG	1712
461	SerIleGluAlaPheArgValGlyTrpGlyAlaLeuGlnTyrGluAspAsnValThrAsn	480
1713	ACCATTTGACAGTTTCGACAGGAGTGGGGTCCCATTTATTATGCTGAG-----TCTAGC	1766
481	ProGluAspMetArgProTyrCysTrpHisTyrProProArgGlnCysGlyValValSer	500
1767	AGATCAGACGAGGCGCATATGCTGGGCACTACCCACCTCCACATGTACCATCTGACTCT	1826
501	AlaLysThrValCysGlyProValTyrCysPheThrProSerProValValValGlyThr	520
1827	CGCTCGGAGGTGTGCGGCGCAGTACTGCTTCACCCAAGGCCCTGCTGCTCGTGGGAGCG	1886
521	ThrAspArgLeuGlyAlaProThrTyrThrTrpGlyGluAsnGluThrAspValPheLeu	540
1887	ACCGATCGTTTCGGTGTCCCTACGATATAGATGGGGGGAGAACGACAGCTGCTGCTGTG	1946
541	LeuAsnSerThrArgProProLeuGlySerTrpPheGlyCysThrTrpMetAsnSerSer	560
1947	CTCAACAACACGGCGCGCGCAAGGCAACTGGTTGCGCTGCACATGATGAATGACACC	2006
561	GlyTyrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer	580
2007	GGGTTCCACAAGACATGTGGGGGGCCCGCTGTATACATCGGGGGGCTCGGCAAC-----	2060
581	ThrAspLeuLeuCysProThrAspCysPheArgLysHisProAspThrThrTyrLeuLys	600
2061	AACACCCCTTGACCTGCCCCACGGACTGCTTCGGGAAGCACCCCGAGGCTACCTACACA	2120
601	CysGlySerGlyProTrpLeuThrProArgCysLeuIleAspTyrProTyrArgLeuTrp	620
2121	TGTGTTTCGGGGCCTTGCTGACCTAGGTGCATGTTGTGATATCCATACAGGCTCTGG	2180
621	HisTyrProCysThrValAsnTyrThrIlePheLysIleArgMetTyrValGlyGlyVal	640
2181	CATTACCCTTCGACCTGTTTAACCTTTACCATCTTCAAGGTTAGATGATATGTGGGGGGGTG	2240
641	GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp	660
2241	GAGCACAGGCTCAATGCTGCATGCAATTTGGACCCGAGGAGAGCGTGTGACTTTGAGGAC	2300
661	ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGluTrpAlaIleLeuPro	680
2301	AGGGATAGCGGGAGCTCAGCCCGCTGCTGTCCTCAACACAGAGTGGCAGGTACTTGCCC	2360
681	CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisLeuHisGlnAsnIle	700
2361	TGTTCCCTTCAACCCCTACCAGCTCTGTCCACTGCGCTTGATTCACTCCATCAGAATCATC	2420
701	ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrLysTyrIleValArgTrp	720
2421	GTGACGTCGCAATACCTATACGATACGGGTACGGGTGTCTCTCTTTCGATCAATAGG	2480

Copied from 09540843 on 05/19/2004

Db	4641		ACAGACCTCTGATACGGGCTATACGGGCGACTTTGACTCAGTGATCGACTGTAAACACA	4700	Db	5721	ATCACAGCCCGCTCACCAACCAAGTACCCTCTCTGTTTAAACATCTTGGGGGGTGGTG	5780
Qy	1461		AlaValThrGlnValAlaAspPheSerLeuAspProThrPheThrIleThrGlnIle	1480	Qy	1821	AlaSerGlnIleAlaProProAlaGlyAlaThrGlyPheValValSerGlyLeuValGly	1840
Db	4701		TGTGTACCCAGACAGTCGACTTCAGCTTGATCCCACTTACCACTTGAGACGAGACC	4760	Db	5781	GCTGCCAACTGCCCGCCCGAGCGCGCTTCGGCTTTCGTGGGGCCCGGCATCGCGGT	5840
Qy	1481		ValProGlnAspAlaValSerArgSerGlnArgGlyArgThrGlyArgGlyArgLeu	1500	Qy	1841	AlaAlaValGlySerIleGlyLeuGlyLysValLeuValAspIleLeuAlaGlyTyrGly	1860
Db	4761		GTGCCTCAAGACGAGTGTCTGCGCTCGCAGCGCGGGTAGGACTGCGAGGGTAGGAGA	4820	Db	5841	CGCGCTGTTGGCAGCATAGCCCTTGGGAAGTGTCTGTGGACATCTCGCGGGTTATGGA	5900
Qy	1501		GlyIleTyrArgTyrValSerThrGlyLysArgAlaSerGlyMetPheAspSerValVal	1520	Qy	1861	AlaGlyIleSerGlyAlaLeuValAlaPheLysIleMetSerGlyLysProSerMet	1880
Db	4821		GGCATCTACAGGTTGTGACTCCGGGAGAACGGCCCTCGGGGCATGTTTCGATTCCTCGGTC	4880	Db	5901	GCAGGAGTGGCGCGCGCTCGTGGCTTTAAGGTTCATGAGCGCGGAGATGCCCTCCACC	5960
Qy	1521		LeuCysGluCysTyrAspAlaGlyAlaAlaTyrTyrGluLeuThrProSerGluThrThr	1540	Qy	1881	GluAspValValAsnLeuLeuProGlyIleLeuSerProGlyAlaLeuValValGlyVal	1900
Db	4881		CTGTGTGAGTGTATGACGGGGCTGTGCTTGTGTACAGCTCACCCGGCGGAGACCTCG	4940	Db	5961	GAGGACCTGGTCAATCTACTTCTGTCATCTCTCTCTGGCGCCCTGGTCTCGGGGTC	6020
Qy	1541		ValArgLeuArgAlaTyrPheAsnThrProGlyLeuProValCysGlnAspHisLeuGlu	1560	Qy	1901	IleCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnThrMet	1920
Db	4941		GTTAGGTTGGGGCTTACCTGAACACACACAGGGTGGCCGTTTGGCAGGACCACTGGAG	5000	Db	6021	GTGTGTGCAACAATCTGGTTCGACACAGTGGGTCCGGGAGAGGGGGCTGTGCAGTGGATG	6080
Qy	1561		PheTyrGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThr	1580	Qy	1921	AsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValAlaProThrHisTyrValThr	1940
Db	5001		TTCCTGGAGAGTGTCTTACAGGCTCACCATATAGATGACACTTCTTGTCTCCAGACC	5060	Db	6081	AAACCGCTGATAGCTTCGCTCGGGGTAAATCATGTTTCCCGCCACGCACTATGTCCCT	6140
Qy	1581		LysGlnSerGlyGluAsnPheAlaTyrLeuThrAlaTyrGlnAlaThrValCysAlaArg	1600	Qy	1941	GluSerAspAlaSerGlnArgValThrGlnLeuLeuGlySerLeuThrIleThrSerLeu	1960
Db	5061		AAGCAGCAGGAGCAAACTTCCCTTACCTGGTATACCAACGACCGGTGTGGCGCAGG	5120	Db	6141	GAGAGCGACCGCGCGGTGTACTCAGATCTCTCCAGCTTACCACCTACCTACAGCTG	6200
Qy	1601		AlaLysAlaProProSerTyrAspValMetTyrLysCysLeuThrArgLeuLysPro	1620	Qy	1961	LeuArgArgLeuHisAsnTyrIleThrGluAspCysProIleProCysGlyGlySerTyr	1980
Db	5121		GCTCAGGCCCACTCCATCATGGATCAATGTGGAGTGTCTATACCGCTTGAACCT	5180	Db	6201	CTGAAAAGGCTCCACAGTGGATTAATGAAGACTGCTCCACACCGTGTCCGGCTCGTGG	6260
Qy	1621		ThrLeuValGlyProThrProLeuLeuTyrArgLeuGlySerValThrAsnGluValThr	1640	Qy	1981	LeuArgAspValTyrAspTyrValCysThrIleLeuThrAspPheLysAsnTyrLeuThr	2000
Db	5181		ACGCTGCAGCGGCCAACACCTTGTCTGACAGGCTGGGAGCCGTCAGATGAGGTCAACC	5240	Db	6261	CTAAGGGATGTTGGGACTGGATATGACGGTGTGACTGACTTCAAGACCTGGCTCCAG	6320
Qy	1641		LeuThrHisProValThrLysTyrIleAlaThrCysMetGlnAlaAspLeuGluValMet	1660	Qy	2001	SerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTyrLys	2020
Db	5241		CTCACCACCCCAACCAATATCATGGCATGTCATGTCGGCTGACCTGGAGGTGTC	5300	Db	6321	TCCAGCTCTCGCGAGCTACTTGGAGTCCCTTTTCTCGGCCACCGGGTACAAG	6380
Qy	1661		ThrSerThrTyrValLeuAlaGlyValLeuAlaAlaValAlaTyrCysLeuAla	1680	Qy	2021	GlyValTyrAlaGlyThrGlyIleMetThrThrArgCysProCysGlyAlaAsnIleSer	2040
Db	5301		ACTAGCACCTGGTGTGGGGGAGTCTCTGACGCTCTGCGCGGTATTCCTGACA	5360	Db	6381	GGAGTCTCGCGGGAGACGGCATCATGCAACACCTGCCCATGTGGAGCACAGATCACC	6440
Qy	1681		ThrGlyCysValCysIleIleGlyArgLeuHisIleAsnGlnArgAlaValAlaPro	1700	Qy	2041	GlyAsnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTyr	2060
Db	5361		ACAGGCACTGTGGTCAATGTGGGTAGGATTTCTGTCCGGGAGCGCGCATTTGTTCCC	5420	Db	6441	GGACATGTCAAAACCGTTTCCATGAGGATCGTCGGGCTTAAGACCTGACAGCAACAGTGG	6500
Qy	1701		AspLysGluValLeuTyrGluAlaPheAspGluMetGluCysAlaSerArgAlaAla	1720	Qy	2061	GlnGlyThrPheProIleAsnCysTyrThrGluGlyGlnCysValProLysProAlaPro	2080
Db	5421		GACAGGAGGTTCTTACACGAGGATTCGATGAATGGAAGAGTCCGCTCGCACCTCCCT	5480	Db	6501	CATGGAACATCTCCCACTAACGCATACACACGGGCCCCCTGCACACCTCTCCAGCGCCA	6560
Qy	1721		LeuIleGluGlyGlnArgIleAlaGluMetLeuLysSerLysIleGlnGlyLeu	1740	Qy	2081	AsnPheLysValAlaIleTyrArgValAlaAlaSerGluTyrAlaGluValThrGlnHis	2100
Db	5481		TACATCAGCAGGAGATGAGCTCGCCGAGCAATTCAGCAGAGAGCGCTCGGCTTACTG	5540	Db	6561	AACTATTCTAGCGCGCTGGGGGTGGCGGTGAGGAGTACCTGAGGAGTACGCGCGGTG	6620
Qy	1741		GlnGlnAlaSerLysGlnAlaGlnAspIleGlnProThrValGlnAlaSerTyrProLys	1760	Qy	2101	GlySerTyrHisTyrIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeu	2120
Db	5541		CAAAACGCCAACCAACGAGGAGGCTGCTCCGCTGGGTGGAGTCCAAGTGGCGAGCC	5600	Db	6621	GGGGATTCTCCTACGTAGCGGCGATGACACTGACCAACGTAAGTAAAGTCCCATGCCAGGTT	6680
Qy	1761		ValGluGlnPheTyrAlaLysHisMetTyrAsnPheIleSerGlyIleGlnTyrLeuAla	1780	Qy	2121	ProSerProGluPhePheSerTyrValAspGlyValGlnIleHisArgPheAlaProThr	2140
Db	5601		CTTGAGACATTCGGGGAGAGCACATGTGGAATTCATCAGCGGATACAGTACTTAGCA	5660	Db	6681	CGCGCTCTGAAATCTCTCGGAGGTGGAGCGAGTTCGCTGCACAGGTACGCTCCGGCG	6740
Qy	1781		GlyLeuSerThrLeuProGlyAsnProAlaValAlaSerMetMetAlaPheSerAlaAla	1800	Qy	2141	ProLysProPhePheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal	2160
Db	5661		GGCTTATCCACTCTGCTGGGAACCCCGCAATAGCATCATTTGATGGCATTCACAGCCTCT	5720	Db	6741	TGCAGCCCTCTCTCAGGAGGAGGTATCATTTCCAGTCTGGGCTCAACCAATACCTGGTT	6800
Qy	1801		LeuThrSerProLeuSerThrSerThrIleLeuLeuAsnIleLeuGlyGlyTyrLeu	1820	Qy	2161	GlySerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThr	2180
					Db	6801	GGGTCAACAGTACCATGCGAGCCCGAACCGGATGTAGCAGTGTCTACTTCCATGCTCACC	6860

QY	2181	AspProSerHisIleThrAlaGluThrAlaAlaArgArgLeuAlaArgGlySerProPro	2200	AlaCysGlnLeuThrProProHisSerAlaArgSerLysThrGlyPheGlyAlaLysGlu	2549
Db	6861	GACCCCTCCACATCACAGCAGAAACGGCTAAGGTAGGTGGCCAGGGGGTCTCCCCC	6920	GCCTGCAAGCTGACGCCCCCACTTCGGCCAAATCAAGTTTGGCTATGGGGCAAGGAC	7910
QY	2201	SerGluAlaSerSerAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThr	2220	ValArgSerLeuSerGlyArgAlaValAenHisIleLysSerValTrpLysAspLeuLeu	2569
Db	6921	TCCTTGGCCAGCTCTTCAGCTAGCCAGTGTCTGCGCTTCTTGAAGGCGACATGCACT	6980	GTCCCGAATCTTCCAGCAAGCCCGTTAAACCACTCCCTGCTGGAAGGACTTGGCTG	7970
QY	2221	ThrHisClyLysAlaTyrAspValAspMetValAspAlaSerLeuPhe	2236	GluAspSerGluThrProThrLeuMetAlaLysAsnGluValPheCysVal	2589
Db	6981	ACCCACCATCTCTCCGAGCGCTGACCTCATCGAGGCCAACCTCTCTGTCGGCGAGGAG	7040	GAAGACACTGTGACCAATTCACACCACATCATGGCAAAATAGAGGTTTCTGTGTC	8030
QY	2237	MetGlyGlyAspValThrArgIleGluSerGlySerLysValValValLeuAspSerLeu	2256	AspProThrLysGlyGlyLysAlaAlaArgLeuLeuValTrpProAspLeuGlyVal	2609
Db	7041	ATGGCGGGGAACATCACCCCGTGGAGTCCGAGAACAGGTGTGTCTGAGTCTTTC	7100	CAACCAGAGAAAGGAGCGTAAAGCCAGCCCGCTTATCGTATTCACCATCTGGGAGTC	8090
QY	2257	AspProMetValGluGluArgSerAspLeuGluProSerIleProSerGluThrMetLeu	2276	ArgValCysGluLysMetAlaLeuThrAspIleThrGlnLysLeuProGlnAlaValMet	2629
Db	7101	GACCCGCTTCGAGCGGAGGAGGATGAGAGGAAGTATCCGTTCCGGCGGAGATCTTCGG	7160	CGTGTATGCGAGAAGATGCCCTCTATGATGTGGTCTCCACCCTTCTCAGGTCGTGATG	8150
QY	2277	ProLysLysArgPheProAlaLeuProAlaTrpAlaArgProAspTyrAsnProPro	2296	GlyAlaSerThrGlyPheGlnTyrSerProAlaGlnArgValGluPheLeuLysAla	2649
Db	7161	AAATCCAAAGATTCCCGCAGCGATGCCCATCTGGCGCGCCCGGATTACAAACCTCCA	7220	GGCTCTCATACGATTCCAGTACTCTCTGGGCGAGGAGTCAAGTTCTGTGGAATACC	8210
QY	2297	LeuValGluSerTrpLysArgProAspTyrGlnProAlaThrValAlaGlyCysAlaLeu	2316	TrpAlaGluLysLysAspProMetGlyPheSerTyrAspThrArgCysPheAspSerThr	2669
Db	7221	CTGTGAGTCTCGGAAGGACCCGAGTACGTCCCTCCGGTGTGACGGGTGCCCGTTG	7280	TGGAAATCAAGAAACCCCATGGCTTTTTCATATGACACTCGCTGTTCGACTCAACG	8270
QY	2317	ProProProArgLysThrProThrProProProArgArgArgThrValGlyLeuSer	2336	ValThrGluArgAspIleArgThrGluGluSerIleThrArgAlaCysSerLeuProGlu	2689
Db	7281	CCACTATCAAGGCCCTCCAAATACCACCTCCACGGAGAAAGAGGCGGTGTCCTAACA	7340	GTCCCGAGAACACATCCGTGTGAGGAGTCAATTTACCAATGTGTGACTTGGCCCCC	8330
QY	2337	GluAspSerIleGlyAspAlaLeuGlnGlnLeuAlaIleLysSerPheGlyGlnProPro	2356	GluAlaHisThrAlaIleHisSerLeuThrGluArgLeuTyrValGlyLysProMetPhe	2709
Db	7341	GAGTCTCTCGGTCTTCTGCTTAGCGGAGCTCGCTACTAAGACCTTCGGC	7391	GAAGCAGACAGCCCAATAAAATCGCTCACAGAGCGCTTTATTCGGGGCTCTCTGACT	8390
QY	2357	ProSerGlyAspSerGlyLeuSerThrGlyValAlaGlyAlaAlaAspSerGlySerGlnThr	2376	AsnSerLysGlyGlnThrCysGlyTyrArgArgCysArgAlaSerGlyValLeuThrThr	2729
Db	7392	AGCTCCGAATCATCGGCCCTCGACAGCGGACCGGACCGGACCGGACCGGACCG	7430	AAATCAAGAGGCGAGAACTGCGGTTATCGCGGTGCGCGCGAGCGCGTGTGACGACT	8450
QY	2377	ProProProAspGluLeuAlaLeuAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla	2390	SerMetGlyAsnThrIleThrCysTyrValLysAlaLeuAlaAlaCysLysAlaAlaGly	2749
Db	7431	GCCCTTCTGACACGAGCTCCGACGAGCGGTGACAAAGGATCCGACGTTGAGTCTGACTCC	7490	AGCTGCGGTAAACCCCTCATCTTAAAGAGCCTCTGCAGCCTGTGCAGCTGCGAAG	8510
QY	2391	SerMetProProLeuGluGlyGluLeuGlyAspProAspLeuGluProGluGlnValGlu	2410	IleAlaProThrMetLeuValCysGlyAspAspLeuValIleSerGluSerGln	2769
Db	7491	TCCATGCCCCCTTGAAGGGGAACCGGGGACCCCGATCTC	7532	CTCCAGGACTGACGAGTGTCTGTGAACGAGAGCAGCTCTGCTGTATCTGTGAACGCGG	8570
QY	2411	ProGlnProProProGlnGlyGlyValAlaAlaProGlySerAspSerGlySerTrpSer	2430	GlyThrGluGluAspGluArgAsnLeuArgAlaPheThrGluAlaMetThrArgTyrSer	2789
Db	7533	AGTGACGGGTCTTGCT	7550	GGAAACCAAGAGGACCGCGGAGCCTACGAGTCTTCCAGGAGGCTATGACTAGTACTCC	8630
QY	2431	ThrCysSerGluGluAspAspSerValValCysCysSerMetSerTyrSerTrpThr	2449	AlaProProGlyAspProProArgProGluTyrAspLeuLeuIleThrSerCysSer	2809
Db	7551	ACCGTAGCGAGGAGTAGTAGGTGTCTGTCTGCTCAATGTCTTACACATGAGACA	7610	GGCCCCCGGGGACCGGCCCAACAGAAATACGACTTGGAGCTGATAACATCATGTTC	8690
QY	2450	GlyAlaLeuIleThrProCysSerProGluGluGluLysLeuProIleAsnProLeuSer	2469	SerAsnValSerValAlaLeuGlyProGlnGlyArgArgTyrTyrLeuThrArgAsp	2829
Db	7611	GGCGCTTGATACGCCATCGCTCGGAGGAAGCAAGCTGCCATCAACGCGGTGAGC	7670	TCCAATGTGTGCGTCCGCCACCGATGCGATCAGGCAAAAGGTTGACTACCTCACCCGATG	8750
QY	2470	AsnSerLeuLeuArgTyrHisAsnLysValTyrCysThrThrThrThrLysSerAlaSerLeu	2489	ProThrThrProIleAlaArgAlaAlaTrpGluThrValArgHisSerProValAsnSer	2849
Db	7671	AACTCTTTGTCGCCCAACCAATCATGTTATGCCCAACATCTCCGACGAGGCGCTG	7730	CCCAACCAACCCCTAGCAGGCTGCTGGGAGACAGCTAGACACACTCCAGTTAACTCC	8810
QY	2490	ArgAlaLysLysValThrPheAspArgMetGlnValLeuAspSerTyrTyrAspSerVal	2509	TrpLeuGlyAsnIleLeuGlnTyrAlaProThrIleTrpAlaArgMetValLeuMetThr	2869
Db	7731	CGGAGAGAGAGGTTCACCTTTGACAGTCAAGTCTTGGACGACACTACCGGAGCTG	7790	TGGCTAGGCAACATTAATATGATGCGCCCATTTCTGGGCAAGGATGATTTCTGATGACT	8870
QY	2510	LeuLysAspIleLysLeuAlaAlaSerLysValThrAlaArgLeuLeuThrMetGlu	2529	HisPhePheSerIleLeuMetAlaGlnAspThrLeuAspGlnAsnLeuAsnPheGluMet	2889
Db	7791	CTCAAGGAGATGAAGGGGAGCGCTCACAGTTAAGGCTAACTCTTATCCGTAGAGGAA	7850	CACCTTCTTCCATCTTCTAGCGAGGAGCAACTTGAAGAACCCCTGGACTCCAGACT	8930
QY				TyrGlyAlaValTyrSerValSerProLeuAspLeuProAlaIleIleGluArgLeuHis	2909

Db 8931 TAGGGGCGCTTACTCCATTCAGCCACTTGACCTACCTCAGATCATTTGAACGACTCCAT 8990
 Qy 2910 GlyLeuAspAlaPheSerLeuHisThrTyrThrProHisGluLeuThrArgValAlaSer 2929
 Db 8991 GGCCTTAGCCATTTTCACTCCATAGTATCTCCAGGTGAGATCAATAGGTGGCTTCA 9050
 Qy 2930 AlaLeuArgHisLeuGlyAlaProProLeuArgAlaTPrlySerArgAlaVal 2949
 Db 9051 TGCCCTCAGGAACCTTGGGGTACCACCTTCGGAGTCTGGAGACATCGGGCCAGGAGGTC 9110
 Qy 2950 ArgAlaSerLeuLeuSerArgGlyGlyArgAlaValCysGlyArgTyrLeuPheAsn 2969
 Db 9111 CGCGCTAGGCTACTGCTCCAGGGAGGAGGGCGCCACTTGTGGCAATACCTCTTCAAC 9170
 Qy 2970 TrpAlaValLysThrLysLeuLysLeuThrProLeuProGluAlaArgLeuLeuAspLeu 2989
 Db 9171 TGCGCAGTAACAACTTAACCTCACTCCACTCCGCTGGCTCCGCTGGAGCTTG 9230
 Qy 2990 SerSerTrpPheThrValGlyAlaGlyGlyGlyAspIleTyrHisSerValSerArgAla 3009
 Db 9231 TCCGGCTGGTTCGTTGCTGTATCAGCGGGGAGACATATATCACAGCCTGTCTCGTGCC 9290
 Qy 3010 ArgProArgLeuLeuPheGlyLeuLeuLeuPheValGlyValGlyLeuPheLeu 3029
 Db 9291 CGACCCCGTGGTTCATGCTGTGCTACTCTCTTCTGTAGGGGTAGGCATCTACCTG 9350
 Qy 3030 LeuProAlaArg 3033
 Db 9351 CTCCCAACGA 9362

RESULT 12
 US-09-315-850-1
 : Sequence 1, Application US/09315850
 : Patent No. 6217872
 : GENERAL INFORMATION:
 : APPLICANT: OKAYAMA, Hiroto
 : APPLICANT: FURE, Isao
 : APPLICANT: MORI, Chisato
 : APPLICANT: TAKAMIZAWA, Akahisa
 : APPLICANT: YOSHIDA, Iwao
 : TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
 : TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
 : NUMBER OF SEQUENCES: 50
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Armstrong, Westerman, Hattori, McLeLland &
 : ADDRESSEE: Naughton
 : STREET: 1725 K St. N.W. Suite 1000
 : CITY: Washington
 : STATE: D.C.
 : COUNTRY: U.S.A.
 : ZIP: 20006
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
 : SOFTWARE: ASCII
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/315,850
 : FILING DATE:
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US/08/904,686
 : FILING DATE: 01-AUG-1997
 : APPLICATION NUMBER: US 08/324,977
 : FILING DATE: 18-OCT-1994
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: JP 2-167466
 : FILING DATE: 25-JUN-1990
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: JP 2-230921
 : FILING DATE: 31-AUG-1990
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: JP 2-305605

: FILING DATE: 09-NOV-1990
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/099,706
 : FILING DATE: 30-JUL-1993
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 07/769,996
 : FILING DATE: 02-OCT-1991
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 07/635,451
 : FILING DATE: 28-DEC-1990
 : ATTORNEY/AGENT INFORMATION:
 : NAME: McLeLland, Le-Nhung
 : REGISTRATION NUMBER: 31,541
 : REFERENCE/DOCKET NUMBER: 900703G
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (202) 659-2930
 : TELEFAX: (202) 887-0357
 : INFORMATION FOR SEQ ID NO: 1:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 9416 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : MOLECULE TYPE: cDNA from genomic RNA
 : FEATURE:
 : NAME/KEY: CDS
 : LOCATION: 333..9362
 : US-09-315-850-1
 : Alignment Scores:
 : Pred. No.: 0 Length: 9416
 : Score: 11984.50 Matches: 2175
 : Percent Similarity: 83.02% Conservativity: 352
 : Best Local Similarity: 71.45% Mismatches: 472
 : Query Match: 74.35% Indels: 45
 : DB: 3 Gaps: 8
 : US-09-980-559-2 (1-3033) x US-09-315-850-1 (1-9416)
 : Qy 1 MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGln 20
 : Db 333 ATGAGCAGCAATCTTAACCTCAAGAAAACCAACGTAACCAACCCGCCGCCACAG 392
 : Qy 21 AspValLysPheProGlyGlyGlyGlnIleValGlyGlyValLysValLysLeuProArg 40
 : Db 393 GACGTCACTTCCCGGGCGGTGTCAGATCGTTGGTGGAGTTTACTGTTGCCGCCAGG 452
 : Qy 41 GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly 60
 : Db 453 GGGCCCGAGGTGGGTGTCGCGCGCCAGGAAGACTTCCGAGCGGTCCGCAACCTCGTGA 512
 : Qy 61 ArgArgGlnProLysProLysAspArgArgSerThrGlyLysSerThrGlyLysProGly 80
 : Db 513 AGCGAGCAACCTTATCCCCAAGCTCGCGCGCCGAGGGCAGGAGCTGAGCGCTCAGCCGGG 572
 : Qy 81 TyrProTrpProLeuTyrGlyAsnGluGlyLeuGlyTyrAlaGlyTyrLeuLeuSerPro 100
 : Db 573 TACCCCTTGGCCTCTCTATGGCAATGAGGGCTTAGGGTGGCAGGATGGCTCCTGTCA 632
 : Qy 101 ArgGlySerArgProSerTrpGlyProAsnAspProArgHisArgSerArgHisValGly 120
 : Db 633 CGCGGCTCCCGGCTAGTTGGGGCGCCACGAGCCCGCGCGGTAGGTCCGGTAAATTTGGGT 692
 : Qy 121 LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProValVal 140
 : Db 693 AAGGTATCATGATACCTCATATGCGGCTTCGCGCATCTCATGGGTACATTCCTCGCTC 752
 : Qy 141 GlyAlaProLeuGlyValAlaArgAlaLeuAlaHisGlyValArgValLeuGluAsp 160
 : Db 753 GGGGCCCCCTGGGGGGCGCTCCAGGGCCCTGGACATGGTGTCCGGGTTCCTGGAGGAC 812
 : Qy 161 GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerPheIlePheLeuAla 180

813 GCGGTGAATATGCAACAGGGAATCTGCCGGTTGCTCTTTTCTATCTCTCTTGGCT 872
 181 LeuLeuSerCysIleThrThrProValSerAlaAlaGluValIleSerThrGly 200
 873 CTGCTGCTCCCTGAGACACCCAGCTTCGCTTACGAGTGCACACGCTGCCGGATA 932
 201 TyrMetValThrAsnAspCysThrAsnAspSerIleThrTriPThrGlnLeuAlaVal 220
 933 TATCATGTACGACGACTGCTGCTCAACGCAAGCATGTGTATGAGGCGGACGTGATC 992
 221 LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTrpIle 240
 993 ATGCATACTCTGGGTGCGTGGCTGGTTCGGGAGGCAACTCTCCGCTGCTGGTA 1052
 241 ProValSerProAsnValAlaValGlnAraProGlyAlaLeuThrGlnGlyLeuArgThr 260
 1053 GCGCTACTCCACGCTCGAGCGAGGAGGAGCTCACCATCCCAACGAGCATACGACGC 1112
 261 HistLeuAspMetValValMetSerAlaThrLeuCysSerAlaLeuTyrValGlyAspLeu 280
 1113 CAGTCGATCTGCTGCTGGGGCGCTGCTTCTGTTCCGCTATGTACGTGGGGGACCTC 1172
 281 CysGlyGlyValMetLeuAlaAlaGlnMetPheIleValSerProGlnHisIleTrpPhe 300
 1173 TGGGATCTGTTTCTCTCTCTCAGCTGTTCACCTCTCTCGCTCGCCGCGCATGTGACA 1232
 301 ValGlnAspCysAsnCysSerIleTyrProGlyThrIleThrGlyHisArgMetAlaTrp 320
 1233 TTACAGGACTGTAACTGCTCAATTTATCCCGGCGCATGTGCGGTACCCGATGTGGTGG 1292
 321 AspMetMetAsnTrpSerProThrAlaThrMetIleLeuAlaTyrAlaMetArgVal 340
 1293 GACATGATGATGAATCTGCTGCCCAACAGCCCTAGTGTGTGCGAGTACTCCGGATC 1352
 341 ProGluValIleLeuAspIleLeuSerGlyAlaHisTrpGlyValMetPheGlyLeuAla 360
 1353 CCACAGCGCTGCTGGGACATGCTGGCGGGGCGCCACTGGGAGTCTCTGGCGGCTTGGC 1412
 361 TyrPheSerMetGlnGlyAlaTrpAlaIleValValIleLeuLeuAlaAlaGly 380
 1413 TACTATTCCATGGCGGGGAACTGGGCTTAAGGTTCTGATTGTGATCTTTTGTGGTC 1472
 381 ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr 400
 1473 GTTACGGGGATACCACTGACAGGGGGGGCGGCAACCAACCAACCAACAGGCTCGT 1532
 401 SerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTrp 420
 1533 TCCATGTTGCAAGTGGCGCTCTCAGAAATCCAGCTTATAACACCAATGGAGTTGG 1592
 421 HistLeuAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer 440
 1593 CACATCAACAGGACTGCGCTGAATGCAATGACTCTCTCCAGACTGGGTTCTTGGCGCG 1652
 441 LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg 460
 1653 CTGTTTACACATAGTTTCACTGCTCGGGTCCCGAGGCGCATGGCCAGCTGCGCGC 1712
 461 SerIleGluAlaPheArgValGlyTrpGlyAlaLeuGlnTyrGluAspAsnValThrAsn 480
 1713 ACCATTGCAAGTTTCGACCGGATGGGTCCTTACTTATCTGAG-----TCTAGC 1766
 481 ProGluAspMetArgProTyrCysTrpHisTyrProArgGlnCysGlyValValSer 500
 1767 AGATCAGACAGAGGCCATATTTGTCGCACTACCCACTCCCAATGTACCATCGTACCT 1826
 501 AlalysThrValCysGlyProValTyrCysPheThrProSerProValValGlyThr 520
 1827 GCGTCGAGGTTGCGGGCCAGTGTACTGCTTCAACCCCAAGCCCTGCTGCTGGGAGC 1886
 521 ThrAspArgLeuGlyAlaProThrTyrThrTrpGlyGluAsnGluThrAspValPheLeu 540
 1887 ACCGATCGTTTTCGTTGCTCCCTACGTATAGATGGGGGAGAACGAGACTGACGCTGCTG 1946

541 LeuAsnSerThrArgProLeuGlySerTrpPheGlyCysThrTrpMetAsnSerSer 560
 1947 CTCACAACACGCGCGCCGCAAGCACTGGTTTCGCTGCACATGATGAATGACACC 2006
 561 GlyTyrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer 580
 2007 GGGTTTCAACAGACATGTGGGGGCGCCCGTGTAAACATCGGGGGGTGCGCAAC----- 2060
 581 ThrAspLeuLeuCysProThrAspCysPheArgLysHisProAspThrThrTyrLeuLys 600
 2061 AACACCCCTGACCTGCCCGACGCTTCCGGAAGACCCCGAGGCTACCTTACACAAA 2120
 601 CysGlySerGlyProTrpLeuThrProArgCysLeuIleAspTyrProTyrArgLeuTrp 620
 2121 TGTGGTTCGGGCGCTTGGCTGACACTAGTGTGATGGTGTACTATCCATACAGGCTCTGG 2180
 621 HistTyrProCysThrValAsnTyrThrIlePheLysIleArgMetTyrValGlyGlyVal 640
 2181 CATTTACCCCTGACCTGTTAACTTTACCATCTTCAAGGTTAGGATGTATGTGGGGGGTGG 2240
 641 GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp 660
 2241 GAGCACAGGCTCAATGCTGCAATTTGACCCGAGGAGAGCGTTGTGACTGTGGAGGAC 2300
 661 ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGluTrpAlaIleLeuPro 680
 2301 AGGATAGCGCGGAGCTCAGCCCGTCTGCTGTCTACACAGAGTGGCAGGCTACTGCC 2360
 681 CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisLeuHisGlnIle 700
 2361 TGTTCCTTCCACCACTACAGCTCTGCTCACTGCTGCTTGTATTCACCTCCATCAGAACATC 2420
 701 ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrLysTyrIleValArgTrp 720
 2421 GTGACGCTGCAATACCTATACGGTATAGGCTCAGCGGTTGTCTCTTTCGCAATCAATGG 2480
 721 GluTrpValIleLeuLeuPheLeuLeuAlaAspAlaArgValCysAlaCysLeuTrp 740
 2481 GAGTATGCTGCTGCTGCTTCTCTCTAGCGGACGACGCTGCTGCTGCTGCTGCTGCTG 2540
 741 MetLeuIleLeuLeuGlyGlnAlaGluAlaLeuGluLysLeuValIleLeuHisAla 760
 2541 ATGATGCTGCTGATAGCCCGAGCGCGCTTGGAGAACCTGCTGCTCTCAATTCG 2600
 761 AlaSerAlaAlaSerCysAsnGlyPheLeuTyrPheValIlePhePheValAlaAlaTrp 780
 2601 GCGTCTGCTGCGCGCGCACATGCGATCTCTCTCTTGTGTCTTCTGTCGCGCTGG 2660
 781 TyrIleLysGlyArgValValProLeuAlaThrTyrSerLeuThrGlyLeuTrpSerPhe 800
 2661 TATCATCAAGGACGCTGCTCTCTGCGGCGACATATGCTCTTTATGCGCTGTCGCGCTG 2720
 801 SerLeuLeuLeuAlaLeuProGlnGlnAlaTyrAlaTyrAspAlaSerValHisGly 820
 2721 CTCTCTCTCTGCTGCAATTACCCAGCGAGCTTACGCCATGCGAGCGGAGATGGTGCA 2780
 821 GlnIleGlyAlaAlaLeuValMetIleThrLeuPheThrLeuThrProGlyTyrLys 840
 2781 TCGTGGGAGCGCGGTTTTGTGGGTCTGGTACTCTCTGCTTGTGTCTTCTGTCGCGCTGG 2840
 841 ThrLeuLeuSerArgPheLeuTrpLeuCysTyrLeuLeuThrLeuGlyGluAlaMet 860
 2841 GTGTTCTCTGCTAGGCTCATATGTTGTTACAATATTTTACCACGAGCGGAGCGGAC 2900
 861 ValGlnGluTrpAlaProMetGlnValArgGlyGlyArgAspGlyIleIleTrpAla 880
 2901 TTACATGTGTGGATCCCGCCCTCAACGCTCGGGGAGCGCGCATGCTCATCTCTCTC 2960
 881 ValAlaIlePheTyrProGlyValValPheAspIleThrLysTrpLeuAlaValLeu 900
 2961 ATGTCGCGAGTCCATCCAGGCTTAATCTTTGACATCACCACAACTTCTTAATTGCCATCTC 3020

901	GlyProAlaTyLeuLeuLeuysGlyAlaLeuLeuThrArgValProTyrPheValArgAlaHis	920
3021	GGTGGCTCATGTGCTCCAAGCTGGCAATAACACAGAGTGCCTACTTCGTGGCGGCTCAA	3080
921	AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyClyArgTyrValGlnMet	940
3081	GGGCTCATTCATGCATGCATGTTAGTGGGAAGTGCCTGGGGGTCAATTATGTGTCAAATG	3140
941	AlaLeuLeuAlaGlyArgTyrThrGlyTyrThrIleTyrAspHisLeuThrProMet	960
3141	GCCTTCATGAAGCTGGCGCGCTGCAGACGACGTACATTTACACCATCTTACCCCGCTA	3200
961	SerAspTyrAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIleIlePhe	980
3201	CGGGATTTGGCCACGGCGCGGCTTACGAGACCTTTCGGTGGCAGTGGAGCCGCTGCTTC	3260
981	SerProMetGluLysLysValIleValTyrGlyAlaGluThraAlaCysGlyAspIle	1000
3261	TCCGACATGGAGCACCAAGATCATCACTGGGAGACAGACACCGCGCGTGTGGGACATC	3320
1001	LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp	1020
3321	ATCTTGGGTCTCGCGCTCTCCGCCGAGAGGANAAGAGATACTCTCTGGCGCGCGCGAT	3380
1021	GlyTyrThrSerLysGlyTrpSerLeuLeuAlaProIleThrAlaTyrAlaGlnThr	1040
3381	AGTCTTGAAGCGCGGGGTTCGACTCTCTCGGCCCATCACGGCTTACTCCCAACAGACG	3440
1041	ArgGlyLeuLeuGlyThrIleValSerMetThrGlyArgAspLysThrGluGlnAla	1060
3441	CGGGCGCTACTTGGTTGGCATCATCACTAGCCCTTACAGCGCGGCAACAGACGCTCGAG	3500
1061	GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly	1080
3501	GGAGAGTTCAAGTGTGTTTCCACCGCAACACATCTCTCTGGCAGCTTCGCTCAACGGC	3560
1081	ValLeuTyrThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro	1100
3561	GTGTGTTGGACCGTTTACCATGTGTGCTCAAGACCTTAGCCGCGCAAAAGGGGCCA	3620
1101	ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTyrProSerProGly	1120
3621	ATCACCCAGATGTACATAATGTGGACACGAGACCTCTCGGTGGCCCAAGCCCCCGGG	3680
1121	ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArgAsn	1140
3681	GGCGGTCTCTGACACCATGACCTGTGGAGCTCAGACTTTACTTGTGTACAGACAT	3740
1141	AlaAspValIleProAlaArgArgGlyAspLysArgGlyAlaLeuLeuSerProArg	1160
3741	GCTGAGCTCATTCGGGTGCGCGGGGGCGAGTAGGGGAGCCTGCTCTCCCCCAGG	3800
1161	ProLeuSerThrLeuLysGlySerSerGlyProValLeuCysProArgGlyHisAla	1180
3801	CCGTGCTCTACTTGAAGGGCTCTTCGGGTGGTCCACTGCTCTGCCCTTCGGCGCACGT	3860
1181	ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle	1200
3861	GTGGGATCTTCGGGCTCCCGTATGCACCCCGGGGGTTCGGAAGCGGTGGACTTTGG	3920
1201	ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro	1220
3921	CCCGTAGAGTCCATGGAACACTACTATGCGGTCTCCGGTCTTCACGGACAACTCATCCCC	3980
1221	ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLys	1240
3981	CGGGCGGTACCCGAGTCATTTCAAGTGGGCCACCTACACGCTCCCACTGGCAGCGCAAG	4040
1241	SerThrLysValProValAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnPro	1260
4041	AGTACTAAAGTCCCGCTGCATATGCAGCCCAAGGGTCAAGGTGCTCTGCTCTCAATCG	4100
1261	SerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerLysAlaHisGlyIleAsnPro	1280

[illegible]

2001	SerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTyrLys	2020
6321	TCCAAGCTCTCCCGCAGACTACCTGGAGTCCCTTTTTTCTCGTCCAAACGGCGGTACAG	6380
2021	GlyValTrrAlaGlyThrGlyLeuMetThrThrArgCysProCysGlyValaAsnIleSer	2040
6381	GGAGTCTGGCGGGAGACGGCATCATGCAAAACCACTGCCCCATGTGGAGCACAGATCAC	6440
2041	GlyAsnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTrrp	2060
6441	GGACATGTCAAAAACGGTTCATGAGGATCGTCGGCGCTAAGACCTGCAGCAACACGTGG	6500
2061	GlnGlyThrPheProIleAsnCysTyrThrGluGlyGlnCysValProLysProAlaPro	2080
6501	CATGNAACATTCCTCCCATACAGCATACACACGGGCCCTGCACACCTCTCCAGGCGCA	6560
2081	AsnPheLysValAlaIleTrrpArgValAlaAlaSerGluTyrAlaGluValThrGlnHis	2100
6561	AACATTATTGAGCGCTGTGGCGGTGGCCGTGAGGAGTAGTACGTGGAGGTCAACGGGGTG	6620
2101	GlySerTyrHisTyrIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeu	2120
6621	GGGGAATTTCCTACGTACCGCGCATGACCACTGCAACCGTAAGTGGCCATGCCAGGTT	6680
2121	ProSerProGluPhePheSerTrrpValAspGlyValGlnIleHisArgPheAlaProThr	2140
6681	CGGCTCTCGAATCTCTCGAGGTGACGGAGTGGCGTTCACAGGTACGCTCGGCG	6740
2141	ProLysProPheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal	2160
6741	TGCAGCTCTCTACGGGAGGAGTTACATTTCCAGGTCCGGCTCAACCAATACCTGGTT	6800
2161	GlySerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThr	2180
6801	GGGTACACAGCTACCATGACGCGCCGACCGGATGTAGCAGTGCTCACTTCATGTCTACC	6860
2181	AspProSerHisIleThrAlaGluThrAlaAlaArgArgLeuAlaArgGlySerProPro	2200
6861	GACCCCTCCACATCACAGCAAAACGGCTAAGCGTAGTGTGGCCAGGGGTCTCCCC	6920
2201	SerGluAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThr	2220
6921	TCTTTGGCGCAGCTCTTTCAGCTAGCAGTTGTCTGCGCTTCCTTGAAGCGCATGCAC	6980
2221	ThrHisGlyLysAlaTyrAspValAspMetValAspAlaAsnLeuPhe	2236
6981	ACCCACCATGTCTCTCCGACGCTGACCTCATCGAGGCCAACCTCTCTGTGGCGGAGGAG	7040
2237	MetGlyGlyAspValThrArgIleGluSerGlySerLysValValValLeuAspSerLeu	2256
7041	ATGGCGGGGAACATCACCGCGGTGGAGTCGGAGAACAAAGGTGGTAGTCTCTGACACTTTC	7100
2257	AspProMetValGluGluArgSerAspLeuGluProSerIleProSerGluTyrMetLeu	2276
7101	GACCCGCTTCGAGCGGAGGAGGATGAGGAGGAAAGTATCGTTCCGGCGGAGATCTGCGG	7160
2277	ProLysLysArgPheProProAlaLeuProAlaTrrpAlaArgProAspTyrAsnProPro	2296
7161	AAATCCAGAAGTTCCCCGACAGCGATGCCATCTGGCGCGCCCGGATTTACAACCTCCA	7220
2297	LeuValGluSerTrrpLysArgProAspTyrGlnProAlaThrValAlaGlyCysAlaLeu	2316
7221	CTGTTTAGAGTCTCTGGAAGACCCGACACTACGTCCTCCGGTGGTGCACGGGTGCCCGTTG	7280
2317	ProProProAlaGlyThrProThrProProArgArgArgArgThrValGlyLeuSer	2336
7281	CCACCTATCAAGGCCCTCCATACCACTCCACGGAGAAAGAGGACGGTTGTCTCTAACA	7340
2337	GluAspSerIleGlyAspAlaLeuGlnLeuAlaIleLysSerPheGlyGlnProPro	2356
7341	GAGTCTCCGTCTCTCTGCTTAGCGAGCTCGCTACTAAGACCTTCGGC	7391

QY 2357 ProSerGlyAspSerGlyLeuSerThrGlyAlaAlaAspSerGlySerGlnThr 2376
DB 7392 -----AGCTCGAATCATCGCGCTCGACAGCGGCACCGGACC 7430
QY 2377 Pro---ProAspGluLeuAla-----LeuSerGluThrGlySerIleSer 2390
DB 7431 GCCCTTCCTGACAGCGCTCCGACGCGGTGACAAAGGATCCGACGTTGAGTCTGACTCC 7490
QY 2391 SerMetProProLeuGluGlyLeuLeuGlyAspProAspLeuProGluGlnValGlu 2410
DB 7491 TCCATGCCGCCCTTGAAGGGGAACCGGGGAGCCCGAATC-----7532
QY 2411 ProGlnProProGlnGlyValAlaAlaProGlySerAspSerGlySerTrpSer 2430
DB 7533 -----AGTCAGCGGTCTTGCTCT 7550
QY 2431 ThrCysSerGluGlu---AspAspSerValValCysSerMetSerTrpThr 2449
DB 7551 ACCGTGAGCGAGAGAGCTAGTGAGGATGCTGCTGCTCAATGTCTACACATCGACA 7610
QY 2450 GlyAlaLeuIleThrProCysSerProGluGluGluLeuProIleAsnProLeuSer 2469
DB 7611 GCGCCCTTGATACGCGATCGCTGCGGAGGAAGCAAGCTGCCATCAACGCTTGAGC 7670
QY 2470 AsnSerLeuLeuArgTyrHisAsnLysValTyrCysThrThrThrLysSerAlaSerLeu 2489
DB 7671 AACTCTTTGCTGCGCCACCATACATGTTTATGCCACAAACATCTCGAGCGCAGCCTG 7730
QY 2490 ArgAlaLysValThrPheAspArgMetGlnValLeuAspSerTyrTyrAspSerVal 2509
DB 7731 CGCAGAGAAGGTCACTTTGACAGCTGCAAGTCTTGACAGCACTACCGGGACGTG 7790
QY 2510 LeuLysAspIleLysLeuAlaAlaSerLysValThrAlaArgLeuLeuThrMetGluGlu 2529
DB 7791 CTACAGAGATGAAGCGGAGGCGTCCACAGTTTAAAGCTAAACTCTATCCGTAGAGGA 7850
QY 2530 AlaCysGlnLeuThrProProHisSerAlaArgSerLysTyrGlyPheGlyAlaLysGlu 2549
DB 7851 GCCTGAAGCTGACGCGCCCAACATTCGCGCAATCCAAAGTTTGGCTATGGGCAAGGAC 7910
QY 2550 ValArgSerLeuSerGlyArgAlaValAsnHisIleLysSerValTrpLysAspLeuLeu 2569
DB 7911 GTCCGAACCTATCCAGCAGGCGGTAAACACATCCACTCCGTGTGGAAGGACTTGTCTG 7970
QY 2570 GluAspSerGluThrProIleProThrIleMetAlaLysAsnGluValPheCysVal 2589
DB 7971 GAAGACACTGTGACCAATTTGACACCATCATGGCAAAATGAGGTTTTCTGTGTC 8030
QY 2590 AspProThrLysGlyGlyLysLysAlaAlaArgLeuIleValTyrProAspLeuGlyVal 2609
DB 8031 CAACAGAGAAAGAGCGCGTAAGCCAGCGCGCTTATCGTATTCACAGATCTGGGAGTC 8090
QY 2610 ArgValCysGluLysMetAlaLeuTyrAspIleThrGlnLysLeuProGlnAlaValMet 2629
DB 8091 CGTGTATGCGAGAAGATGGCCCTCTATGATGTGCTCTCCACCCCTTCTCAGGTCGTGATG 8150
QY 2630 GlyAlaSerTyrGlyPheGlnTyrSerProAlaGlnArgValGluPheLeuLysAla 2649
DB 8151 GGCCTCATACCGATTCAGATGACTCTCTGGGCGAGGTCGAGTCTCTGGTAATACC 8210
QY 2650 TrpAlaGluLysLysAspProMetGlyPheSerTyrAspThrArgCysPheAspSerThr 2669
DB 8211 TGGAAATCAAGAAACCCATGGGCTTTTCATATGACATCGCTGTTTTCGACTCAACG 8270
QY 2670 ValThrGluArgAspIleArgThrGluGluSerIleTyrArgAlaCysSerLeuProGlu 2689
DB 8271 GTCAACGAGAACGACATCCGTGTGTGAGGAGTCAATTTACCAATGTTGTGACTTGGCCCC 8330
QY 2690 GluAlaHisThrAlaIleHisSerLeuThrGluArgLeuTyrValGlyGlyProMetPhe 2709
DB 8331 GAAGCCAGACAGGCCCAATAAATCGCTCACAGAGCGGCTTTATATCGGGGTCCTCTGACT 8390
QY 2710 AsnSerLysGlyGlnThrCysGlyTyrArgArgCysArgAlaSerGlyValLeuThrThr 2729

DB 8391 AATTCAAAAGGCGAGAACTGCGGTATTCGCGGTGCGCGCGAGCGCGTGTCTGACGACT 8450
QY 2730 SerMetGlyAsnThrIleThrCysTyrValLysAlaLeuAlaAlaCysLysAlaAlaGly 2749
DB 8451 AGTCGCGGTAAACCCCTCATGTTTACTTTGAAGGCTCTGCGAGCCCTGTGAGTCTGGAAG 8510
QY 2750 IleIleAlaProThrMetLeuValCysGlyAspAspLeuValIleSerGluSerGln 2769
DB 8511 CTCAGGAGTGCACGATGCTCGTGAACGAGAGCACTCTGCTGTTATCTGTGAAGCGCG 8570
QY 2770 GlyThrGluGluAspGluArgAsnLeuArgAlaPheThrGluAlaMetThrArgTyrSer 2789
DB 8571 GGAAACCAAGAGAGCGCGGAGCTTACGAGTCTTCAGCGAGCTATGACTAGGTACTCC 8630
QY 2790 AlaProProGlyAspProProArgProGluTyrAspLeuLeuIleThrSerCysSer 2809
DB 8631 GCGCCCCCGGGAGCCCGCCCAACAGAAATACGACTTGGAGCTGATAACATCATGTTC 8690
QY 2810 SerAsnValSerValAlaLeuGlyProGlnGlyArgArgTyrTyrLeuThrArgAsp 2829
DB 8691 TCCATGTGCTCGTCCGCCCATGATGCATCAGGCAAAAGGGTGTACTACCTCACCGTAT 8750
QY 2830 ProThrThrProIleAlaArgAlaAlaTrpGluThrValArgHisSerProValAsnSer 2849
DB 8751 CCGCACCCCGCTAGCAGCGCTGCTGGGAGACAGCTAGACACACTCCAGTTAACTCC 8810
QY 2850 TrpLeuGlyAsnIleGlnTyrAlaProThrIleTrpAlaArgMetValLeuMetThr 2869
DB 8811 TGGCTAGGCAACATATTATGATGCGCCCTTTGTGGGCAAGGATGATTTCTGATGACT 8870
QY 2870 HisPhePheSerIleLeuMetAlaGlnAspThrLeuAspGlnAsnLeuAsnPheGluMet 2889
DB 8871 CACTTCTTCTCATCTTCTAGCGGAGGCACTTGAAAAAGCCCTGGACTGCCAGATC 8930
QY 2890 TyrGlyAlaValTyrSerValSerProLeuAspLeuProAlaIleIleGluArgLeuHis 2909
DB 8931 TAGCGGCGCTGTACTTCCATTCAGCCACTTGACCTACCTCAGATCATTTGAACGACTCCAT 8990
QY 2910 GlyLeuAspAlaPheSerLeuHisThrTyrThrProHisGluLeuThrArgValAlaSer 2929
DB 8991 GCGCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTCGATCAATAGGGTGGCTTCA 9050
QY 2930 AlaLeuArgLysLeuGlyAlaProProLeuArgAlaAlaTrpLysSerArgAlaArgAlaVal 2949
DB 9051 TGCTCAGGAACACTTGGGGTACACCTTGCGAGTCTGAGACATCGGCCAGGAGGCTC 9110
QY 2950 ArgAlaSerLeuIleSerArgGlyArgAlaAlaValCysGlyArgTyrLeuPheAsn 2969
DB 9111 CGCGCTTAGGCTACTGTCCTCAGGAGGCGCGCCACTTGTGGCAATACCTCTTCAAC 9170
QY 2970 TrpAlaValLysThrLysLeuLysLeuThrProLeuProGluAlaArgLeuAspLeu 2989
DB 9171 TGGGAGTAGTAAAAACCAAACTTAAACTCACTCAATCCCGGCTCGCGCTGGAGCTTG 9230
QY 2990 SerSerTrpPheThrValGlyAlaGlyGlyLysAspIleTyrHisSerValSerArgAla 3009
DB 9231 TCGGCTGTGTTGTTGCTGTTTACAGCGGGGAGACATATATCAGACCTGTCTCTGTC 9290
QY 3010 ArgProArgLeuLeuLeuPheGlyLeuLeuLeuPheValGlyValGlyLeuPheLeu 3029
DB 9291 CGACCCCGTGTGTTTCATGCTGCTTACTCTTCTTCTAGGGGTAGGCATCTTACTCTG 9350
QY 3030 LeuProAlaArg 3033
DB 9351 CTCCCCAACCGA 9362

RESULT 13

US-08-823-895A-27
; Sequence 27, Application US/08823895A
; Patent No. 6433159
; GENERAL INFORMATION:
; APPLICANT: Kevin P. Anderson


```
Qy 481 ProGluAspMetArgProTyrCysTyrHisTyrProProArgGlnCysGlyValValSer 500
Db 1767 AGATCAGACAGAGGCCATATTCCTGGCACTACCCACCTCCACCAATGTACCATCGTACCT 1826
Qy 501 AlaLysThrValCysGlyProValTyrCysPheThrProSerProValValGlyThr 520
Db 1827 GCCTCGAGAGTGTGCGCGCCAGTGTATCTGCTTACCCCAAGCCCTGCTGCTGCGGGGACG 1886
Qy 521 ThrAspArgLeuGlyAlaProThrTyrThrTyrGlyGluAsnGluThrAspValPheLeu 540
Db 1887 ACCGATCGTTTGGTGTCCCTAGTATAGATGGGGGAGAACGAGACTGACGTGCTGCTG 1946
Qy 541 LeuAsnSerThrArgProProLeuGlySerTyrPheGlyCysThrTyrMetAsnSerSer 560
Db 1947 CTCAACAACAGCGCGCGCGCAAGGCAACTGCTCGGCTGCACATGGATGAATGACAC 2006
Qy 561 GlyTyrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer 580
Db 2007 GGGTTCAACCAAGACATGTGGGGGGCCCCCGTGTAAACATCGGGGGGTCGCGCAAC----- 2060
Qy 581 ThrAspLeuLeuCysProThrAspCysPheArgLysHisProAspThrThrTyrLeuLys 600
Db 2061 AACACCTGACCTGCCCCAGGACTGCTTCCGGAAGCACCCTCGAGGCTACCTACACAAA 2120
Qy 601 CysGlySerGlyProTyrLeuThrProArgCysLeuIleAspTyrProTyrArgLeuTyr 620
Db 2121 TGTGGTTCGGGGCTTGGCTGACACCTAGTGTGATGGTTGACTATCCATACAGGCTCTGG 2180
Qy 621 HisTyrProCysThrValAsnTyrThrIlePheLysIleArgMetTyrValGlyGlyVal 640
Db 2181 CATTTACCCCTGCACTGTAACTTTACCATCTTCAAGGTTAGGATGTATGTGGGGGGGTG 2240
Qy 641 GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp 660
Db 2241 GAGCAGGCTCAATGCTGATCAATGACCCGAGGAGAGGTTGTGACTTTGGAGGAC 2300
Qy 661 ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGluTyrAlaIleLeuPro 680
Db 2301 AGCGATAGCGCGAGCTAGCCCGCTGCTGCTGTCTACCAACAGAGTGGCAGGTACTGCC 2360
Qy 681 CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisLeuHisGlnAsnIle 700
Db 2361 TGTTCCTTCCACCACTACAGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2420
Qy 701 ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrLysTyrIleValArgTyr 720
Db 2421 GTGGACGTGCAATACCTATACGGTATAGGCTCAGCGGTTGTCTCTTTTGCATCAATGG 2480
Qy 721 GluTyrValIleLeuLeuPheLeuLeuAlaAspAlaArgValCysAlaCysLeuTyr 740
Db 2481 GAGTATGCTCTGTTTCTTCTCTCTAGCGGACGACGTGCTGTGCTGCTGCTGCTGCTGCT 2540
Qy 741 MetLeuIleLeuLeuGlyGlnAlaGluAlaLeuLeuLysLeuValIleLeuHisAla 760
Db 2541 ATGATCTGCTGATGCCAGCGGAGCGCGCTTGGAGAACCTGCTGCTGCTGCTGCTGCTGCT 2600
Qy 761 AlaSerAlaAspCysAsnGlyPheLeuTyrPheValIlePhePheValAlaAlaTyr 780
Db 2601 GCGTCTGTGCGCGCGCACATGATGCTCTCTCTCTGTTGTGTTCTGTCGCGCGCTG 2660
Qy 781 TyrIleLysGlyArgValValProLeuAlaThrTyrSerLeuThrGlyLeuTyrSerPhe 800
Db 2661 TACATCAAGGAGGCTGCTGCTGCTGCGGCGACATATGCTCTTTATGGCGTGTGCGCGCTG 2720
Qy 801 SerLeuLeuLeuAlaLeuProGlnAlaTyrAlaTyrAspAlaSerValHisGly 820
Db 2721 CTCCTGCTCTGCTGATACCAACCGCGAGCTTACGCCATGGACCGGAGATGGGTGCA 2780
Qy 821 GlnIleGlyAlaAlaLeuLeuMetIleThrLeuPheThrLeuThrProGlyTyrLys 840
Db 2781 TCGTGGGAGGCGCGGTTTTTGTGGTCTGGTACTCTCTGACTTGTGTCACCATACTACAAG 2840
Qy 841 ThrLeuLeuSerArgPheLeuTyrTyrPheLeuCysTyrLeuLeuThrLeuGlyGluAlaMet 860
Db 2841 GTGTTCTCGCTAGGCTCATATGGTGTGTACAAATATTTTACCAACAGAGCGGAGCGGAC 2900
Qy 861 ValGlnGluTyrAlaProProMetGlnValArgGlyGlyArgAspGlyIleIleTyrAla 880
Db 2901 TTACATGTGTGATCCCCCTCCCAACGCTCGGGGAGCGCGATGCAATCATCTCTCTC 2960
Qy 881 ValAlaIlePheTyrProGlyValValPheAspIleThrLysTyrPheLeuAlaValLeu 900
Db 2961 ATGTGGCAGTCTCATCCAGAGCTAATCTTTGACATCACCACAACTCTTAATGCGCATCTC 3020
Qy 901 GlyProAlaTyrLeuLeuLysGlyAlaLeuThrArgValProTyrPheValArgAlaHis 920
Db 3021 GGTCCGCTCATGTGCTCCCAAGCTGGCATAACACAGAGTCCCGTACTTCTGTGGCGCTCAA 3080
Qy 921 AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyGlyArgTyrValGlnMet 940
Db 3081 GGGCTCATTTCAATGCAATGCAATGTTAGTGGCGAAGTCTGCTGGGGGTCTATTATGTCCAATG 3140
Qy 941 AlaLeuLeuAlaLeuGlyArgTyrThrGlyThrTyrIleTyrAspHisLeuThrProMet 960
Db 3141 GCCTTCATGACCTGGCGCGCTGACAGCAGCATATTTACACCATCTTACCCCGCTA 3200
Qy 961 SerAspTyrAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIleIlePhe 980
Db 3201 CGGGATGGCCACGCGCGGCTTACGAGACCTTGGCTGGCAGTGGAGCCCGTCTCTTC 3260
Qy 981 SerProMetGluLysIleValIleValTyrGlyAlaGluThrAlaAlaCysGlyAspIle 1000
Db 3261 TCCGATCGAGACCAAGATCATCTGCGGAGCAGACACCGCGCGTGTGGGGACATC 3320
Qy 1001 LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp 1020
Db 3321 ATCTTGGGTCTCCCGCTCTCCCGCGAGGGAAGAGATACTCTCTGGCGCGCGCGAT 3380
Qy 1021 GlyTyrThrSerLysGlyTyrSerLeuLeuAlaProIleThrAlaTyrAlaGlnThr 1040
Db 3381 AGTCTTGAAGCGCGGGTGGGACTCTCTCGCGCCATCACGGCTTACTCCCAACAGACG 3440
Qy 1041 ArgGlyLeuLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGluGlnAla 1060
Db 3441 CGGGCGCTACTTGGTGTGATCATCATCTAGCTTACAGCGCGGCAACAGAGCTCGAG 3500
Qy 1061 GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly 1080
Db 3501 GGAGAGTTCAGTGGTGTTCACCGCAACAACTCTCTCTGGCAGCTGCGTCAACGGC 3560
Qy 1081 ValLeuTyrThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGlyPro 1100
Db 3561 GTGTGTGGACCGTGTACCATGTGCTGGCTCAAGACCTTAGCGCGCAAGGGGCCA 3620
Qy 1101 ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTyrProSerProGly 1120
Db 3621 ATACCCAGATGTACACTAATGTGACAGGAGCTCTGCTGGCTGGCGCCAGGCCCCCGG 3680
Qy 1121 ThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArgAsn 1140
Db 3681 GCGGCTCTTTCACACCATGACCTGTGGCAGCTTACCTTTTGTGTGTCAGAGACAT 3740
Qy 1141 AlaAspValIleProAlaArgArgGlyAspLysArgGlyValAlaLeuLeuSerProArg 1160
Db 3741 GCTGACGTATTCGCTGCGCGCGCGGCGAGACAGTAGGGGAGCTGTCTCTCCCCCAG 3800
Qy 1161 ProLeuSerThrLeuLysGlySerSerGlyGlyProValLeuLeuCysProArgGlyHisAla 1180
Db 3801 CCTGTCTCTTCTTGAAGGGCTCTTGGGTGTGCTCACTGCTGCTGCTGCTGCTGCTGCTG 3860
Qy 1181 ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle 1200
Db 3861 GTGGGACATCTTCGGGCTGCGGTATGACCCCGGGGGTGTGGAAGCGGTGACATTGTG 3920
Qy 1201 ProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro 1220
```


Copied from 09540843 on 05/19/2004

```
QY 1941 GluSerAspAlaSerGlnArgValThrGlnLeuLeuGlySerLeuThrIleThrSerLeu 1960
Db 6141 GAGAGCGAGCGCGAGCGGTGTACTCAGATCTCTCCAGCCTTACCATCATCTCAGCTG 6200
QY 1961 LeuArgArgLeuHisIleThrGlnLeuGlySerProIleProCysGlyGlySerTrp 1980
Db 6201 CTGAAGAGGCTCCAGCAGTGGATTAAATGAAGACTGCTCCACACCGGTTCGCGCTCGTG 6260
QY 1981 LeuArgAspValTrpAspTrpValCysThrIleLeuThrAspPheLysIleThrLeuThr 2000
Db 6261 CTNAGGATGTTTGGAGCTGGATATGACGGTGTGACTGACTTCAAGACCTGGCTCCAG 6320
QY 2001 SerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTrpLys 2020
Db 6321 TCCAGCTCTCGCGCAGCTACCTGGAGTCCCTTTTCTCGTGCCCAACGCGGTACAAG 6380
QY 2021 GlyValTrpAlaGlyThrGlyIleMetThrThrArgCysProCysGlyAlaIleSer 2040
Db 6381 GGAGCTCTGGCGGAGCAGCGCATCATGCAACCACTGCCCATGTGGAGCACAGATCAC 6440
QY 2041 GlyAsnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetIleTrp 2060
Db 6441 GGACATGTCAAAACCGTTCATGAGATCGTGGGCGCTTAAGACCTGCAGCAACAGTGG 6500
QY 2061 GlnGlyThrPheProIleAsnCysTrpThrGluGlyGlnCysValProLysProAlaPro 2080
Db 6501 CATGGAACATCTCCCATCAACGATACACACCGGCGCTGCACACCTCTCCAGGCGCA 6560
QY 2081 AsnPheLysValAlaIleTrpArgValAlaIleSerGluTrpAlaGluValThrGlnHis 2100
Db 6561 AACTATTCTAGGCGCTGTGGCGGTGGCGGTGAGGAGTACGTGGAGGTCAACGCGGTG 6620
QY 2101 GlySerTrpHisTrpIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeu 2120
Db 6621 GGGGATTTCCATCTACGTACGCGCATGACCATGCAACGTAAGTAAAGTGGCCATGCCAGTT 6680
QY 2121 ProSerProGluPhePheSerTrpValAspGlyValGlnIleHisArgPheAlaProThr 2140
Db 6681 CCGGCTCCTGMAATCTCTCGAGGTGGAGCGGTGCGGTGACAGGTACGCTCGGCG 6740
QY 2141 ProLysProPhePheArgAspGluValSerPheCysValGlyLeuAsnSerPheValVal 2160
Db 6741 TGCAGGCGCTCTCTACGGGAGGAGTTACATTCAGGTGCGGTCAACCAATACCTGGTT 6800
QY 2161 GlySerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThr 2180
Db 6801 GGTCTACAGCTACCATGCGAGCGCGAACCGGATGTAGCAGTGTCTCATCTCTCACC 6860
QY 2181 AspProSerHisIleThrAlaGluThrAlaAlaArgArgLeuAlaArgGlySerProPro 2200
Db 6861 GACCCCTCCCATCATCACAGCAAGAACGGCTAGGTTGGCCAGGGGGTCTCCCCC 6920
QY 2201 SerGluAlaSerSerAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThr 2220
Db 6921 TCCTGGCCAGCTCTTTCAGCTAGCCAGTTGTCTGCGCTTCTTGAAGCGCATGCATC 6980
QY 2221 ThrHisGlyLysAlaTrpAspValAspMetValAspAlaAsnLeuPhe 2236
Db 6981 ACCACCATGTCTCTCCGACCGCTGACCTCATCGAGCGCAACCTCTCTGTGGCGCAGGAG 7040
QY 2237 MetGlyGlyAspValThrArgIleGluSerGlySerLysValValValLeuAspSerLeu 2256
Db 7041 ATGGCGCGGACATCACCGCGTGGAGTCGGAGAACAGGTGTGTCTCTGGACTCTTTC 7100
QY 2257 AspProMetValGluGluArgSerAspLeuGluProSerIleProSerGluTrpMetLeu 2276
Db 7101 GACCCGCTTCGAGCGGAGGAGGATGAGAGGATATCGTTCCGCGGAGATCTCTGCGG 7160
QY 2277 ProLysLysArgPheProAlaLeuProAlaIleTrpAlaArgProAspTrpAsnProPro 2296
Db 7161 AAATCCAAAGAGTTCCCGCAGCGATGCCCATCTGGCGCGCGCCGATTAACAACCTCCA 7220
QY 2297 LeuValGluSerTrpLysArgProAspTrpGlnProAlaThrValAlaGlyCysAlaLeu 2316
Db 7221 CTGTTAGAGTCTCGAAGGACCGGACTAGCTCCCTCCGCTGCTGCACGGGTGCCGTTG 7280
QY 2317 ProProProArgLysThrProThrProProProArgArgArgThrValGlyLeuSer 2336
Db 7281 CCACCTATCAAGGCCCTCCATACCACTCCACGAGAAAGAGACGGTGTCTCTTAACA 7340
QY 2337 GluAspSerIleGlyAspAlaLeuGlnLeuAlaIleLysSerPheGlyGlnProPro 2356
Db 7341 GAGTCTCTCGCTCTTCTGCTTACGGAGCTCGCTACTAGACCTTCGCG 7391
QY 2357 ProSerGlyAspSerGlyLeuSerThrGlyAlaGlyAlaAlaAspSerGlySerGlnThr 2376
Db 7392 AGCTCCGAATCATCGCGCTCCACAGCGGACCGGAC 7430
QY 2377 ProProAspGluLeuAla 2390
Db 7431 GCCCTTCTGACCGGCTCCGACGCGGTGCAAGAGATCCGAGTTGAGTCTGACTCC 7490
QY 2391 SerMetProProLeuGluGlyGluLeuGlyAspProAspLeuGluProGluGlnValGlu 2410
Db 7491 TCCATGCCCGCTTGGAGGGGACCGGGGACCCCGATCTC 7532
QY 2411 ProGlnProProGlnGlyGlyValAlaAlaProGlySerAspSerGlySerTrpSer 2430
Db 7533 AGTGACGGGTCTTGGTCT 7550
QY 2431 ThrCysSerGluGlu 2449
Db 7551 ACCGTGAGCGAGAAAGTGTGAGGATGTCTGCTCAATGTCTTACATCATGGACA 7610
QY 2450 GlyAlaLeuIleThrProCysSerProGluGluGluLeuLeuProIleAsnProLeuSer 2469
Db 7611 GCGGCTTGTATCACCCTGCTCGGAGGAGAACAGCTGCCATCAACCGTTGAGC 7670
QY 2470 AsnSerLeuLeuArgTrpHisAsnLysValTrpCysThrThrLysSerAlaSerLeu 2489
Db 7671 AACTCTTGTGCGCCACCATAAATGTTTATGCCACAACATCTCGCAGCGCAGGCTG 7730
QY 2490 ArgAlaLysLysValThrPheAspArgMetGlnValLeuAspSerTrpTrpAspSerVal 2509
Db 7731 CGGCAGAGAAGTCACTTTGACAGACTGCAAGTCTTGAGCAGCACCTACCGGACGCTG 7790
QY 2510 LeuLysAspIleLysLeuAlaAlaSerLysValThrAlaArgLeuLeuThrMetGluGlu 2529
Db 7791 CTCAGGAGATGAGCGAGCGTCCACAGTTAAGCTTAAGTCTTATCTCTAGAGGAA 7850
QY 2530 AlaCysGlnLeuThrProHisSerAlaArgSerLysTrpGlyPheGlyAlaLysGlu 2549
Db 7851 GCCTGCAAGCTGACCGCCCAATTCGCGCAATCCAAAGTTTGGCTATGGGCAAGGAC 7910
QY 2550 ValArgSerLeuSerGlyArgAlaValAsnHisIleLysSerValTrpLysAspLeuLeu 2569
Db 7911 GTCCGAACCTATCCAGCAAGCGCTTAACCACTCACTCCGCTGGAAGGACTTGTCTG 7970
QY 2570 GluAspSerGluThrProIleProThrIleMetAlaLysAsnGluValPheCysVal 2589
Db 7971 GAAGACACTGTGACCAATTTGACACCACTCATGGCAAAATATGAGTTTCTGTGTC 8030
QY 2590 AspProThrLysGlyLysLysAlaAlaArgLeuIleValTrpAspLeuGlyVal 2609
Db 8031 CAACAGAGAAAGGAGCGGTAAAGCGCCCTTATCTGATTTCCAGATCTCGGAGTCT 8090
QY 2610 ArgValCysGluLysMetAlaLeuTrpAspIleThrGlnLysLeuProGlnAlaValMet 2629
Db 8091 CGTGATGCGAAGATGCCCTCTATGATGTGGTCTCCACCTTCTCCTCAGGTCGTGATG 8150
QY 2630 GlyAlaSerTrpGlyPheGlnTrpSerProAlaGlnArgValGluPheLeuLysAla 2649
Db 8151 GGCTCTCATACGGATTCAGTACTCTCTGGGAGGAGTCTGAGTCTCTGTGTAATACC 8210
QY 2650 TrpAlaGluLysAspProMetGlyPheSerTrpAspThrArgCysPheAspSerThr 2669
```

8211 TGAATCAAGAAACCCATCGGCTTTTCATATGACACTGCTGTTTCGACTCAACG 8270
 2670 ValThrGluArgAspIleArgThrGluGluSerIleTyrArgAlaCysSerLeuProGlu 2689
 8271 GTCAACGAGAACGACATCCGTTGTGAGAGTCAATTTACCAATGTTGTGACTTGGCCCGC 8330
 2690 GluAlaHisThrAlaIleHisSerLeuThrGluArgLeuTyrValGlyGlyProMetPhe 2709
 8331 GAAGCCAGACAGGCCCAATAAATCGCTCACAGAGCGGCTTTATATCGGGGCTCCTGACT 8390
 2710 AsnSerLysGlyGlnThrCysGlyTyrArgAlaSerGlyValLeuThrThr 2729
 8391 AATTCAAAGGAGACACTGCGGTTATCGCGGTGCGCGAGCGGCTGCTGACGACT 8450
 2730 SerMetGlyAsnThrIleThrCysTyrValIleAlaLeuAlaCysLysAlaAlaGly 2749
 8451 AGCTGGGTTAACACCTCACATGTTACTTGAAGGCTCTGACGCTGTGAGCTGCGAAG 8510
 2750 IleIleAlaProThrMetLeuValCysGlyAspAspLeuValIleSerGluSerGln 2769
 8511 CTCAGGAGACTGCACGATGCTCGTGAACGGAGACGCTCGTCTTATCTGTGAAGCGCG 8570
 2770 GlyThrGluGluAspGluArgAsnLeuArgAlaPheThrGluAlaMetThrArgTyrSer 2789
 8571 GGAACCCAGAGACGCGGCGAGCTACGAGTCTTCAGGAGGCTATGACTAGTACTCC 8630
 2790 AlaProGlyAspProArgProGluTyrAspLeuLeuIleThrSerCysSer 2809
 8631 GCCCCCCCGGAGCCGCCCAACCAACAGATACGACTTGGAGCTGATAACATCATGTTC 8690
 2810 SerAsnValSerValAlaLeuGlyProGlnGlyArgArgTyrTyrLeuThrArgAsp 2829
 8691 TCCAATGTGCGTGGCCACGATGTCATCAGGCAAGAGGCTGACTACCTCCACCGCTGAT 8750
 2830 ProThrThrProIleAlaArgAlaAlaTyrGluThrValArgHisSerProValAsnSer 2849
 8751 CCACCCACCCCTAGACAGCGGTGCTGGAGACAGCTACACACACTCCAGTTAACTCC 8810
 2850 TrpLeuGlyAsnIleGlnTyrAlaProThrIleTrpAlaArgMetValLeuMetThr 2869
 8811 TGGCTAGGCAACATTTATATGATGCGCCCACTTGTGGCAAGGATGATCTGATGACT 8870
 2870 HisPhePheSerIleLeuMetAlaGlnAspThrLeuAspGlnAsnLeuAsnPheGluMet 2889
 8871 CACTTCTCTCCATCTCTAGCGCAGGAGCACTTGAAGAGCCCTGGACTGCCAGATC 8930
 2890 TyrGlyAlaValTyrSerValSerProLeuAspLeuProAlaIleIleGluArgLeuHis 2909
 8931 TAGCGGCGCTGTACTTCATGAGCCACTTGACCTACCTCAGATCATTTGAACGACTCCAT 8990
 2910 GlyLeuAspAlaPheSerLeuHisTyrThrProHisGluLeuThrArgValAlaSer 2929
 8991 GGCCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGTGGCTTCA 9050
 2930 AlaLeuArgLysLeuGlyAlaProProLeuArgAlaTyrLysSerArgAlaAlaVal 2949
 9051 TGCCTCAGAAACTTGGGGTACCACTTGGAGTCTGGAGACATCGGGCCAGGAGCGTC 9110
 2950 ArgAlaSerLeuIleSerArgGlyGlyArgAlaAlaValCysGlyArgTyrLeuPheAsn 2969
 9111 CGCCTTAGGCTACGTCCCGAGGAGGAGGCGCCCACTTGTGGCAATTAACCTTCTCAAC 9170
 2970 TrpAlaValLysThrLysLeuLysLeuThrProLeuProGluAlaArgLeuLeuAspLeu 2989
 9171 TGGCAGTAAAAACCAAACTTAACTCACTCCATCCCGGTGCGTCCCGCTGGACTTG 9230
 2990 SerSerTrpPheThrValGlyAlaGlyGlyAspIleTyrHisSerValSerArgAla 3009
 9231 TCCGGCTGGTGTGCTGTGTACAGCGGGGAGACATATATCACACGCTGTCTCGGCC 9290
 3010 ArgProArgLeuLeuPheGlyLeuLeuLeuPheValGlyValGlyLeuPheLeu 3029

9291 CGACCCCGTGGTTCATGCTGTGCTACTCTACTCTTCTGTAGGGGTAGCATCTACTG 9350
 3030 LeuProAlaArg 3033
 9351 CTCCTCCCAACCGA 9362
 RESULT 14
 US-08-811-566-1
 ; Sequence 1, Application US/08811566
 ; Patent No. 6127116
 ; GENERAL INFORMATION:
 ; APPLICANT: Rice, Charles et al.
 ; TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C
 ; TITLE OF INVENTION: VIRUS (HCV) AND USES THEREOF
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSES: David A. Jackson, Esq.
 ; STREET: 411 Hackensack Ave, Continental Plaza, 4th
 ; STREET: Floor
 ; CITY: Hackensack
 ; STATE: New Jersey
 ; COUNTRY: USA
 ; ZIP: 07601
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/811,566
 ; FILING DATE: 03-MAR-1997
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Jackson Esq., David A.
 ; REGISTRATION NUMBER: 26,742
 ; REFERENCE/DOCKET NUMBER: 1113-1-006
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 201-487-5800
 ; TELEFAX: 201-343-1684
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 9646 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; US-08-811-566-1
 Alignment Scores:
 Pred. No.: 0 Length: 9646
 Score: 11973.00 Matches: 2175
 Percent Similarity: 83.29% Conservative: 357
 Best Local Similarity: 71.55% Mismatches: 472
 Query Match: 74.27% Indels: 36
 DB: 3 Gaps: 8
 US-09-980-559-2 (1-3033) x US-08-811-566-1 (1-9646)
 QY 1 MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGln 20
 DB 342 ATGAGCAGCAATCTTAACCTCAAGAAAACCAACAGTAACACCGTCCGCCACAG 401
 QY 21 AspValLysPheProGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArg 40
 DB 402 GACGTCAAGTTCCTCCGGTGGCGTCAATCGTGGTGGAGTTTACTTGTTCGCGCGCAGG 461
 QY 41 GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly 60
 DB 462 GCGCCCTAGATTGGGTGTCGCGCAGCAGAGAAAGACTTCGAGCGGTCGCAACCTCGAGGT 521
 QY 61 ArgArgGlnProIleProLysAspArgArgSerThrGlyLysSerTrpGlyLysProGly 80

Db 522 AGACGTCAGCCTATCCCAAGGACGTCGCGCCGAGGAGGACCTGGGCTCAGCCCGG 581
Qy 81 TyrProTyrProLeuTyrGlyAsnGluGlyLeuGlyTyrAlaGlyTyrTrpLeuLeuSerPro 100
Db 582 TACCCCTGGCCCTCTATGGCAATAGAGGTGGGGTGGGGATGGCTCTCTGCTCC 641
Qy 101 ArgGlySerArgProSerTyrGlyProAsnAspProArgHisArgSerArgAsnValGly 120
Db 642 CGTGGCTTCGGCTAGCTAGTGGGCCCCACAGACCCCGCGTAGGTGGGCAATTTGGGT 701
Qy 121 LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProValVal 140
Db 702 AAGGTATCATACCTTACGTGGCTTCGCGACCTCATGGGGTACATACCGCTCGTC 761
Qy 141 GlyAlaProLeuGlyGlyValAlaAArgAlaLeuAlaHisGlyValArgValLeuGluAsp 160
Db 762 GCGCCCTCTTGGAGGCGCTGCAGGCGCTCGCGCATGGCGTTCGGGTTCGGAGAC 821
Qy 161 GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla 180
Db 822 GCGCTGAACCTATGCAACAGGAACTTCCTGGTGTCTTCTCTATCTCTCTCTGCGC 881
Qy 181 LeuLeuSerCysIleThrThrProValSerAlaAlaGluValLysAsnIleSerThrGly 200
Db 882 CTGCTCTTGGCTGACTGTGCGCGCTTCAGCCTACCAAGTGGCAATTCCTCGGGCTT 941
Qy 201 TyrMetValThrAsnAspCysThrAsnAspSerIleThrTrpGlnLeuGlnAlaVal 220
Db 942 TACCATGTCAACCAATGTCCTTAACCTAGTGTGTGTACGAGCGCGCGATGCCATC 1001
Qy 221 LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTrpIle 240
Db 1002 CTGCACACTCGGGGTGTCCCTTGGTTCGGAGSGTAACGCTCGAGGTGTGGTG 1061
Qy 241 ProValSerProAsnValAlaValGlnArgProGlyAlaLeu---ThrGlnGlyLeuArg 259
Db 1062 GCGGTGACCCCAACGCTGGCCACC---AGGGAGCGCAAACTCCCAACAGCAGCTTCGA 1118
Qy 260 ThrHisIleAspMetValValMetSerAlaThrLeuCysSerAlaLeuTyrValGlyAsp 279
Db 1119 CGTCATATGATCTGTGTGGAGCGCCACCCTCTGTCTGGCCCTCTACGTGGGGGAC 1178
Qy 280 LeuCysGlyGlyValMetLeuAlaAlaGlnMetPheIleValSerProGlnHisTrp 299
Db 1179 CTGTGGGTCTGTCTTCTTGTGTGCTCACTGTCTTACCTTCTCTCCAGGCGCACTGG 1238
Qy 300 PheValGlnAspCysAsnCysSerIleTyrProGlyThrIleThrGlyHisArgMetAla 319
Db 1239 ACGACGCAAGACTGCAATTTCTATCTATCCCGGCATATACCGGTTCATCGCATGGCA 1298
Qy 320 TrpAspMetMetMetMetMetMetMetMetMetMetMetMetMetMetMetMetMetMet 339
Db 1299 TGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1358
Qy 340 ValProGluValIleIleAspIleIleSerGlyAlaHisTrpGlyValMetPheGlyLeu 359
Db 1359 ATCCCAACAGCCATCATGACATGATGATGATGATGATGATGATGATGATGATGATG 1418
Qy 360 AlaTyrPheSerMetGlnGlyAlaTyrAlaLysValValIleLeuLeuLeuAlaAla 379
Db 1419 GCGTATTTCTCCATGCTGGGGAACGCGGCAAGGCTCTGTGTGTGTGTGTGTGTGTGTGT 1478
Qy 380 GlyValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeu 399
Db 1479 GCGGTGACGCGGAAACCCACGTCACCGGGGGAAGTGGCGGCGGCGGCGGCGGCGGCTT 1538
Qy 400 ThrSerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySer 419
Db 1539 GTTGGTCTCTTACACGCGCGCAAGCAACATCACTGATCAACCAACCAAGCGAGT 1598
Qy 420 TrpHisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAla 439
Db 2726 TGGTATCTGAAGGGTAGGTGGGTGGCGGAGCGGTCTACGCTTCTACGCGGATGTGGCT 2726

QY 800 PheSerLeuLeuLeuAlaLeuProGlnGlnAlaTyrAlaTyrAspAlaSerValHis 819
 Db 2727 CTCCTCTGCTCTGCTGGGTTGCTCTAGCGGGCATAGCACTGACAGGAGGTGGCC 2786
 QY 820 GlyGlnIleGlyAlaAlaLeuLeuValMetIleThrLeuPheThrLeuThrProGlyTyr 839
 Db 2787 GCCTCGTGGCGCGTGTCTTCTGTCGGGTAAATGGCGCTGACTCTGTCCCATATTAC 2846
 QY 840 LysThrLeuLeuSerArgPheLeuThrProMetGlnValArgGlyGlyArgAspGlyIleIleTyr 859
 Db 2847 AAGCGCTACATCAGCTGGTGTGCTGCTTCAAGCCAGTTTCTTAAAGTCCCTTACTTCTGCGCGGT 2906
 QY 860 MetValGlnGlnTrpAlaProProMetGlnValArgGlyGlyArgAspGlyIleIleTyr 879
 Db 2907 CAATCGACGCTGGTGTGCTGCTTCAAGCCAGTTTCTTAAAGTCCCTTACTTCTGCGCGGT 2966
 QY 880 AlaValAlaIlePheTyrProGlyValValPheAspIleThrLysTrpLeuLeuAlaVal 899
 Db 2967 CTCATGTGTGTGACACCGACTCTGCTATTTGACATCACCAAACTACTCTCTGGCCATC 3026
 QY 900 LeuGlyProAlaTyrLeuLeuGlyGlyAlaLeuThrArgValProTyrPheValArgAla 919
 Db 3027 TTCGGACCCCTTTCGATTTCTTCAAGCCAGTTTCTTAAAGTCCCTTACTTCTGCGCGGT 3086
 QY 920 HisAlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyGlyArgTyrValGln 939
 Db 3087 CAAGGCTTCTTCGGATCTCGCGCTAGCGCGGAAGATAGCGAGGTCAATACGTGCA 3146
 QY 940 MetAlaLeuLeuAlaLeuGlyArgTyrThrTyrIleTyrAspHisLeuThrPro 959
 Db 3147 ATGCCCATCATCAAGTTAGGGCGCTTACTGGCACCTATGTGTATAACCATCTCACCCCT 3206
 QY 960 MetSerAspTrpAlaLeuSerGlyLeuArgAspLeuAlaValAlaValGluProIleIle 979
 Db 3207 CTTTCGAGACTGGCGGCACACCGCTCGAGATCTCGCGCTGCTGGAACAGTCTGTC 3266
 QY 980 PheSerProMetGlnLysValIleValTrpGlyValaGluThrAlaAlaCysGlyAsp 999
 Db 3267 TTTCTCCGAATGGAGACCAAGCTCATACGTGGGGGAGATACCGCGCGTGGTGAC 3326
 QY 1000 IleLeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAla 1019
 Db 3327 ATCATCAACGCTTGGCGCTCTGCGCGTAGGGCCAGGAGATCTGCTTGGCCAGCC 3386
 QY 1020 AspGlyTyrThrSerLysGlyTrpSerLeuLeuAlaProIleThrAlaTyrAlaGlnGln 1039
 Db 3387 GACGGAATGCTCTCAAGGGGTGAGGTCTGCGCGCCATCATCGCGCTACGCCAGCAG 3446
 QY 1040 ThrArgGlyLeuLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGluGln 1059
 Db 3447 ACCAGAGGCTCTCTAGGGTGTATATACACAGGCTGACTGGCGGGACAAAACCAAGTG 3506
 QY 1060 AlaGlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSer 1079
 Db 3507 GAGGTGAGTCCAGATCGTGTCACTGCTACCCNAACTTCTTGGCAACGTGCTCAAT 3566
 QY 1080 GlyValLeuTrpThrValTyrHisGlyAlaGlyAsnLysThrLeuAlaGlySerArgGly 1099
 Db 3567 GGGGTATGCTGGACTGTCTACCAACCGGGCGGGAACGAGGACCATCGCATCACCAAGGT 3626
 QY 1100 ProValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTrpProSerProPro 1119
 Db 3627 CTTGTATCATCAGATGTATACATATGGACCAAGACCTTGTGGCTGGCCGCTCTCTCAA 3686
 QY 1120 GlyThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArg 1139
 Db 3687 GGTTCGCGCTCATTGACACCTGACCTGCGGCTCTCTCGGACCTTTTACCTGGTCAAGG 3746
 QY 1140 AsnAlaAspValIleProAlaArgArgArgGlyAspLysArgGlyAlaLeuLeuSerPro 1159
 Db 3747 CAGCCCATGTCTTCCGTCGCGCGGAGGTGATAGAGGGGTGCTTTCGCCC 3806

QY 1160 ArgProLeuSerThrLeuLysGlySerSerGlyGlyProValLeuCysProArgGlyHis 1179
 Db 3807 CGGCCCATTTCTTACTTGAAGGCTCTCGGGGGTCCGCTTGTGCCCCCGGACAC 3866
 QY 1180 AlaValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPhe 1199
 Db 3867 GCGTGGGCTATTTCAGGCGCGGTGTCACCGTGGAGTGGCTAAAGCGGTGGACTTT 3926
 QY 1200 IleProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThr 1219
 Db 3927 ATCCCTGTGGAGAACCTAGACACCACTAGATCCCCGGTGTTCACGGACAACTCTCT 3986
 QY 1220 ProProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGly 1239
 Db 3987 CCACAGCAGTGTCCCGAGAGCTTCCAGGTGGCCACCTGCATGCTCCACCGGACGGT 4046
 QY 1240 LysSerThrLysValProValAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn 1259
 Db 4047 AAGAGCACCAAGGTCCCGCTGCTAGCAGCGCCAGGGCTACAAGGTGTGGTGTCTCAAC 4106
 QY 1260 ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerLysAlaHisGlyIleAsn 1279
 Db 4107 CCTCTGTGTGTCACCGCTGGGCTTGTGTCTTACATGTCCAGGCCCATGGGGTGTAT 4166
 QY 1280 ProAsnIleArgThrGlyValArgThrValThrThrGlyAlaProIleThrTyrSerThr 1299
 Db 4167 CTTATATCAGGACCGGGGTGAGAACATTAACACTGGCAGCCCATCATCATCCACC 4226
 QY 1300 TyrGlyLysPheLeuAlaAspGlyGlyCysAlaGlyGlyAlaTyrAspIleIleIleCys 1319
 Db 4227 TACGGCAAGTCTCTTCCGACGCGGTCTCAGGAGGTGCTTATGACATAATAATTGT 4286
 QY 1320 AspGluCysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGln 1339
 Db 4287 GAGGAGTGCCTTCCACGATGCCATCATCTTGGGCTCGGCACTGCTCTTGACCAA 4346
 QY 1340 AlaGluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProGlySerVal 1359
 Db 4347 GCAGAGACTCGGGGCGGAGACTGGTGTGCTCGCCACTGCTACCCCTCCGGGCTCGTC 4406
 QY 1360 ThrThrProHisProAsnIleGluGluValAlaLeuGlyGlnGluGlyGluIleProPhe 1379
 Db 4407 ACTGTGTCCCATCTTAAACATCGAGGAGGTGCTCTGTCCACCACCGGAGAGATCCCTTT 4466
 QY 1380 TyrGlyArgAlaIleProLeuSerTyrIleLysGlyGlyArgHisLeuIlePheCysHis 1399
 Db 4467 TACGGCAAGCTATCCCCCTCGAGGTGATCAAGGGGGAGAGACATCTCATCTTCTGCCAC 4526
 QY 1400 SerLysLysCysAspGluLeuAlaAlaLeuArgGlyMetGlyLeuAsnSerVal 1419
 Db 4527 TCANAGNAGTGCAGAGCTCGCGGAGAGCTGGTGGCATCAATGGCGTG 4586
 QY 1420 AlaTyrTyrArgGlyLeuAspValSerValIleProThrGlnGlyAspValValValVal 1439
 Db 4587 GCCTACTACCGGCTTCTGACGTGTCTGTCATCCCGACCGCGGCGATGTGTGTCGTG 4646
 QY 1440 AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn 1459
 Db 4647 TCGACGATGCTCTCATGATGGCTTACCGGCACTTCGACTCTGTGTATGACTGCAAC 4706
 QY 1460 ValAlaValThrGlnValValAspPheSerLeuAspProThrPheThrIleThrGln 1479
 Db 4707 ACCTGTGTCTCAGACAGTTCGATTTTCAGCTTTCAGCTTACCTTACCTTACCTTACCT 4766
 QY 1480 IleValProGlnAspAlaValSerArgSerGlnArgArgGlyThrGlyArgGlyArg 1499
 Db 4767 AGCTCTCCCGAGATGCTGTCTCAGGACTCAACCGCGGGGAGGAGTGCAGCGGGGAG 4826
 QY 1500 LeuGlyIleTyrArgTyrValSerThrGlyGluArgAlaSerGlyMetPheAspSerVal 1519
 Db 4827 CCAGGCTATCTACAGATTTGTGGCACCGGGGGAGGCCCTTCCGGCATGTTTCTGCTCC 4886
 QY 1520 ValLeuCysGluCysTyrAspAlaGlyAlaAlaTrpTyrGluLeuThrProSerGluThr 1539

4887	Db		GTTCCTGTGAGTGCATACGCGGGCTGTGCTTGGTATGAGCTACGCGCCGCGAGACT	4946
1540	Qy		ThrValArgLeuArgAlaTyrPheAenThrProGlyLeuProValCysGlnAspHisLeu	1559
4947	Db		ACAGTTAGGCTACGAGCGTACATGAACACCCGCGGGCTTCCCGTGTGCCAGACCATCTT	5006
1560	Qy		GlupheTrpGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln	1579
5007	Db		GAATTTTGGGAGGGCGCTTTACGGGGCTCACTCATATAGATGCGCCACTTCTATCCAG	5066
1580	Qy		ThrLysGlnSerGlyGluAsnPheAlaTyrLeuThrAlaTyrGlnAlaThrValCysAla	1599
5067	Db		ACAAAGACAGTGGGAGAACTTTCCTTACCTGTGTAGCGTACCAAGCCACCGGTGTGCGT	5126
1600	Qy		ArgAlaLysAlaProProSerTrpAspValMetTrpLysCysLeuThrArgLeuLys	1619
5127	Db		AGGGCTCAAGCCCTCCCCCATCTGGGACACAGATGTGGAAGTGTGTGATCGCCCTTAA	5186
1620	Qy		ProThrLeuValGlyProThrProLeuLeuTyrArgLeuGlySerValThrAsnGluVal	1639
5187	Db		CCACCCCTCCATGGGCCAACACCCCTGCTATACAGACTGGCGCTGTTCAGAATGAAGTC	5246
1640	Qy		ThrLeuThrHisProValThrLysTyrIleAlaThrCysMetGlnAlaAspLeuGluVal	1659
5247	Db		ACCTGACGACCCCAATCACCAATACATCATGACATGTCGTGGCCGACCTTGGAGGTC	5306
1660	Qy		MetThrSerThrTrpValLeuAlaGlyGlyValLeuAlaAlaValAlaAlaTyrCysLeu	1679
5307	Db		GTACAGACACTGGGTGCTGTTGGCGGGCTCTGGCTGCTTGGCGCGGTATTGCCCTG	5366
1680	Qy		AlaThrGlyCysValCysIleIleGlyArgLeuHisIleAsnGlnArgAlaValValAla	1699
5367	Db		TCAACAGGCTGGTGTGCATAGTGGCAGATTGCTTCTCGGGAAGCGGCAATTATA	5426
1700	Qy		ProAspLysGluValLeuTyrGluAlaPheAspGluMetGluGluCysAlaSerArgAla	1719
5427	Db		CCTGCAGCGGAGGTCTCTACGAGGAGTTTCGATGAGATGGAAGAGTGCTCTCAGCACTTA	5486
1720	Qy		AlaLeuIleGluGluGlyGlnArgIleAlaGluMetLeuLysSerLysIleGlnGlyLeu	1739
5487	Db		CCGTACATCGACAGGGAGTATGTCGCTGAGCAGTTCAGCAGAGGCGCTCGGCGCTC	5546
1740	Qy		LeuGlnGlnAlaSerLysGlnAlaGlnAspIleGlnProThrValGlnAlaSerTrpPro	1759
5547	Db		CTGCAGACCGCGTCCCGCCACGACAGAGGTATACACCCCTGCTGTCCAGACCACTGGCAG	5606
1760	Qy		LysValGluGlnPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu	1779
5607	Db		AAACTCGAGGTCTTCTGGGCGAAGCACATGTGGAATTTTCATCAGTGGGATACAATACTTG	5666
1780	Qy		AlaGlyLeuSerThrLeuProGlyAsnProAlaValAlaAlaSerMetMetAlaPheSerAla	1799
5667	Db		CGGGGCGCTGTCAACGCTGCTGGTAAACCCCGCATGTGCTTCAATGTATGCTTTTACAGT	5726
1800	Qy		AlaLeuThrSerProLeuSerThrThrIleLeuLeuAsnIleLeuGlyGlyTrp	1819
5727	Db		GGCGTCAACAGCCCATACCACTGGCCAAACCCCTCTCTTCAACATATTGGGGGGGTGG	5786
1820	Qy		LeuAlaSerGlnIleAlaProProAlaGlyAlaThrGlyPheValValSerGlyLeuVal	1839
5787	Db		GTGGCTGCCAGCTGCCGCCCGCCCGGTGCGCTACCGCTTGTGGCGCTGGCTTAGCT	5846
1840	Qy		GlyAlaAlaValGlySerIleGlyLeuGlyLysValLeuValAspIleLeuAlaGlyTyr	1859
5847	Db		GGCGCGCCCATCGGACGCTGGACTGGGGAAGGTCCTCGTGACATCTTCAGGGTAT	5906
1860	Qy		GlyAlaGlyIleSerGlyAlaLeuValAlaPheLysIleMetSerGlyGluLysProSer	1879
5907	Db		GGCGGGGCGTGGCGGAGCTCTTGTATGATTCAGATCATGAGCGGTGAGGTGCCCTCC	5966
1880	Qy		MetGluAspValValAsnLeuLeuProGlyIleLeuSerProGlyAlaLeuValValGly	1899

5967	ACGGAGGACCCTGGTCAATCTGCTGCCCGCCCACTCTCTGCCTGGAGCCCTTGTTAGTCGGT	6026
1900	VallileCysAlaIalleLeuArgArghHisValGlyProGlyGluGlyAlaValGlnTrp	1919
6027	GTGGTCTGGCGAGCAATACTGCGCGGCACOGTTGGCCCGGCGAGGGGGCAGTGCATGG	6086
1920	MetAenArgLeuIlleAlaPheAlaSerArgGlyAenHisValAlaProThrHisTyVal	1939
6087	ATGAACCGGCTTAATAGCCTTCGCCCTCCCGGGGAACCATGTTTCCCACGCACCTACGCTG	6146
1940	ThrGluSerAepAlaSerGlnArgValThrGlnLeuLeuGlySerLeuThrIleThrSer	1959
6147	CCGAGAGCGATGAGCGCCCGCGCTACTGCCACTACTCAGCAGCCTCACGTGTAACCCAG	6206
1960	LeuLeuArgArgLeuHisAenTripleThrGluAspCysProIleProCysGlyGlySer	1979
6207	CTCCTGAGCGCATGTCATCAGTSGNATAGCTCGGAGTGTACCACTCCATGCTCCGGTTCC	6266
1980	TrpLeuArgAspValTrpAspTrpValCysThrIleLeuThrAspPheLysAenTrpLeu	1999
6267	TGGCTTAGGGACATCTGGGACTGGATATGCGAGGTGCTGAGCGACTTTAAAGACCTGGCTG	6326
2000	ThrSerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTy	2019
6327	AAAGCCCAAGCTCATGCCCCAACATGCTGGGATTCCTTTGTCTGCTGCAGCGCGGGTAT	6386
2020	LysGlyValTrpAlaGlyThrGlyIleMetThrThrArgCysProCysGlyAlaAenIle	2039
6387	AGGGGGGTCTGGCGAGGAGACGGCATATGACACTCTGCTGCCACTGTGGAGCTGAGATC	6446
2040	SerGlyAenValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAenIle	2059
6447	ACTGCACATGTCAAAAACGGGACGATGAGGATCGTCGGTCTCTAGGACCTGCAGGAACATG	6506
2060	TrpGlnGlyThrPheProIleAasnCysTyThrGluGlyGlnCysValProLysProAla	2079
6507	TGGAGTGGGACGTTCCTCCCATTAAGCCCTACACACGCGGCCCTGTACTCCCTCTCTCGG	6566
2080	ProAenPheLysValAlaIleTrpArgValAlaAlaSerGluTyThrAlaGluValThrGln	2099
6567	CCGAACATATAGTTCGCGCTGTGAGGGTGTCTGCAGAGGAATAGCTGGAGNTAAGGCGG	6626
2100	HisGlySerTyHisTyRileThrGlyLeuthrThrAspAenLeuLysValProCysGln	2119
6627	GTGGGGGACTTCCACTACGTATCGGTATGACTACTGACAATCTTAAATGCCCGTGCCAG	6686
2120	LeuProserProGluPhePheSerTrpValAspGlyValGlnIleHisArgPheAlapro	2139
6687	ATCCCATCGCCCGCAATTTTTTCACAGAATTGACGGGGTGCOCCTTACATAGGTTTCGCGCC	6746
2140	ThrProLysProPhePheArgAspGluValSerPheCysValGlyLeuAasnerPheVal	2159
6747	CCTTGCAGCCCTTGCTCGGGAGAGAGGTATCATTCAGATGAGGACTCCAGAGTACCCG	6806
2160	ValGlySerGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeu	2179
6807	GTGGGGTGGCAATTACCTTTCGAGCCCGAAACCGGACGTAGCCGTGTTCAGCTGCCATGCTC	6866
2180	ThrAspProSerHisIleThrAlaGluThrAlalaArgArgLeuAlaAargGlySerPro	2199
6867	ACTGATCCTCCCATATACACAGCAGAGCGCGCCGGAGAAAGGTTTGGCGAGAGGGTCACCC	6926
2200	ProSerGluAlaSerSerAlaserGlnLeuSerAlaProSerLeuArgAlaThrCys	2219
6927	CCTTCTATGGCCAGCTCTCTCGGCCAGCAGCTGTCGCTTCATCTCTCAAGGCAACTTGC	6986
2220	ThrThrHisGlyLysAlaTyAspValAspMetValAspAlaAenLeuPhe-----	2236
6987	ACCGCCAACCATGACTCCCTCGACCGCGAGCTCATAGAGGCTTAACCTCTCTGGGAGGCAG	7046
2237	---MetGlyGlyAspValThrArgIleGluSerGlySerLysValValValLeuAaspSer	2255
7047	GAGATGGCGCGCAACATCACGAGGTTGAGTTCAGAGAACAAAGTGGTANTCTTGACATCC	7106

QY	2256	LeuAspProMetValGluGluArgSerAspLeuGluProSerIleProSerGluTyrMet	2275	QY	2614	LysMetAlaLeuTyrAspIleThrGlnLysLeuProGlnAlaValMetGlyAlaSerTyr	2633
DB	7107	TTCCGATCCGCTTGTGGCAGGAGGATGAGCGGAGGTCTCCGTACCCGAGAAATTCGTG	7166	DB	8115	AGATGGCCCTGTACGACGTGGTTAGCAAGCTCCCCCTGGCGTGTATGGGAAGCTCTAC	8174
QY	2276	LeuProLysLysArgPheProProAlaLeuProAlaTyrAlaArgProAspTyrAsnPro	2295	QY	2634	GlyPheGlnTyrSerProAlaGlnArgValGluPheLeuLeuLysAlaTyrAlaGluLys	2653
DB	7167	CGAAGTCTCGGAGATTCGCGCGGCTGCTGCTGGCGCGCGGACTACACCCC	7226	DB	8175	GGATTCCAATACTCCACGAGCAGCGGGTGAATTCCTCGTCAAGCGTGAAGTCAAG	8234
QY	2296	ProLeuValGluSerTrpLysArgProAspTyrGlnProAlaThrValAlaGlyCysAla	2315	QY	2654	LysAspProMetGlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluArg	2673
DB	7227	CCGCTAGTAGACAGCTGGAAAGACCTGACTACGAACACCTGCTGCTGCTGCTGCTG	7286	DB	8235	AAAGCCCGATGGGTTCGTATGATACCCGCTGTTTGGATCTCCACAGTCACTGAGAGC	8294
QY	2316	LeuProProArgLysThrProThrProProArgArgArgThrValGlyLeu	2335	QY	2674	AspIleArgThrGluGluSerIleTyrArgAlaCysSerLeuProGluGluAlaHisThr	2693
DB	7287	CTACCACTCCACGGTCCCTCTGCTGCTCCGCTCGGAAAGAGGTGCGGTGCTC	7346	DB	8295	GACATCCGTACGAGGAGCAATTTACCAATGTTGTGACTTGACCCCAAGCCCGCGTG	8354
QY	2336	SerGluAspSerIleGlyAspAlaLeuGlnGlnLeuAlaIleLysSerPheGlyGlnPro	2355	QY	2694	AlaIleHisSerLeuThrGluArgLeuTyrValGlyGlyProMetPheAsnSerLysGly	2713
DB	7347	ACCAATCAACCTATCTACTGCTTGGCCGAGCTTGGCCACCAAAAGTTTGGCAGCTCC	7406	DB	8355	GCATCAAGTCCCTCACTGAGAGGCTTTATGTTGGGGCCCTCTTACCAATTCAGGGGG	8414
QY	2356	ProProSerGlyAspSerGlyLeuSerThrGlyAlaGlyAlaAlaAspSerGlySerGln	2375	QY	2714	GlnThrCysGlyTyrArgArgCysArgAlaSerGlyValLeuThrThrSerMetGlyAsn	2733
DB	7407	TCAACTTCCGCAATACCGGCGACATATACGACCAATCTCTGAGCCGCGCTTCTGCG	7466	DB	8415	GAAGTCTGGCTTACCGCAGGTGCGCGGAGCGGCTACTGACAACTGCTGTGTATC	8474
QY	2376	ThrProProAspGluLeuAlaLeuSerGluThrGlySerIleSerSerMetProProLeu	2395	QY	2734	ThrIleThrCysTyrValLysAlaLeuAlaCysLysAlaAlaGlyIleIleAlaPro	2753
DB	7467	TGCCCCCCCCGAC-----TCCGAGTGTAGTCTTATCTTCCATGCCCCCCCCCTG	7514	DB	8475	ACCTCCTGTGTACATCAAGGCCCGGCGAGCTGTCGAGCCGAGGCTCCAGGACTGC	8534
QY	2396	GluGlyLeuLeuGlyAspProAspLeuGluProGluGlnValGluProProProPro	2415	QY	2754	ThrMetLeuValCysGlyAspLeuValValIleSerGluSerGlnGlyThrGluGlu	2773
DB	7515	GAGGGGAGCTGGGATCCGGATCTC-----	7541	DB	8535	ACCATGCTGCTGTGGCGACGACTTAGTGTATCTGTGAAGTGTGGGGTCCAGGAG	8594
QY	2416	GlnGlyValAlaAlaProGlySerAspSerGlySerThrTrpSerThrCysSerGlu---	2434	QY	2774	AspGluArgAsnLeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGly	2793
DB	7542	-----AGCAGCGGGTCATGCTCGACGGTCAGTAGTGGG	7574	DB	8595	GACGGCGGCGCTGAGAGCTTTCAGGAGGCTATGACAGGTACTCCGCCCCCCCGGG	8654
QY	2435	---GluAspAspSerValValCysCysSerMetSerTyrSerThrThrGlyAlaLeuIle	2453	QY	2794	AspProProArgProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSer	2813
DB	7575	GCGACACGGAAGATGTCGTGCTGCTCAATGCTTATTCCTGGACGCGGCACTGCTC	7634	DB	8655	GAACCCCAACACCAACAGTACGCTTGGAGCTTATACATCATGCTCTCTCAACGCTGCA	8714
QY	2454	ThrProCysSerProGluGluGluLysLeuProIleAsnProLeuSerAsnSerLeuLeu	2473	QY	2814	ValAlaLeuGlyProGlnGlyArgArgTyrTyrLeuThrArgAspProThrThrPro	2833
DB	7635	ACCCGTCGCTGGGAGAACAAATACTGCCATCAACGCACTGACCACTCGTTGCTA	7694	DB	8715	GTCCGCCACGACGCGGCTCGAAAGAGGCTCTACTCTTACCGCTGACCCCTACACCCCC	8774
QY	2474	ArgTyrHisAsnLysValTyrCysThrThrThrLysSerAlaSerLeuAlaLysLys	2493	QY	2834	IleAlaArgAlaAlaTyrGluThrValArgHisSerProValAsnSerTrpLeuGlyAsn	2853
DB	7695	CGCCATCACAATCTGGTGTATTCACCACTTCACGCACTGTCGCCAAAGGCGAAGAAA	7754	DB	8775	CTCGGAGAGCGCGGGGAGACAGCAGACACTCCAGTCAATCTCTGGCTGAGCAAC	8834
QY	2494	ValThrPheAspArgMetGlnValLeuAspSerTyrTyrAspSerValLeuLysAspIle	2513	QY	2854	IleIleGlnTyrAlaProThrIleTyrAlaArgMetValLeuMetThrHisPheSer	2873
DB	7755	GTCAATTTGACAGCTGCAAGTCTCGACAGCCATTCACGAGCGTCTCAAGGAGGTC	7814	DB	8835	ATAATCATCTTTGCCCCACACTGTGGCGGAGGATGATGATGATGATGATGATGATGATG	8894
QY	2514	LysLeuAlaAlaSerLysValThrAlaArgLeuLeuThrMetGluGluAlaCysGlnLeu	2533	QY	2874	IleLeuMetAlaGlnAspThrLeuAspGlnAsnLeuAsnPheGluMetTyrGlyAlaVal	2893
DB	7815	AAAGCAGCGGCTCAAAAGTGAAGGCTAACTTGTATCCGTAGGAGGCTTGCAGGCTG	7874	DB	8895	GTCTCATAGCCAGGATCAGCTTGAACAGGCTCTTAACGTGTGAGATCTACGAGGCTGC	8954
QY	2534	ThrProProHisAlaArgSerLysTyrGlyPheGlyAlaLysGluValArgSerLeu	2553	QY	2894	TyrSerValSerProLeuAspLeuProAlaIleIleGluArgLeuHisGlyLeuAspAla	2913
DB	7875	ACGCCCCACATTCAGCCAAATCAAGTTTGGCTATGGGCGAAAGACGCTCCGTTGCCAT	7934	DB	8955	TACTTCATAGAACCACTGGATCTACCTCAATTCATTAAGAGCTCCATGGCTCAGCGCA	9014
QY	2554	SerGlyArgAlaValAsnHisIleLysSerValTrpLysAspLeuLeuAspSerGlu	2573	QY	2914	PheSerLeuHisThrTyrThrProHisGluLeuThrArgValAlaSerAlaLeuArgLys	2933
DB	7935	GCCAGAAAGCGGTAGCCCAATCACTCCGTTGGGAAAGACCTTCTGGGAAGACAGTGA	7994	DB	9015	TTTTTCACTCCAGTTACTCTCCAGGTGAATCAATAGGGTGGCCGATGCCTCAGAAAA	9074
QY	2574	ThrProIleProThrThrIleMetAlaLysAsnGluValPheCysValAspProThrLys	2593	QY	2934	LeuGlyAlaProProLeuArgAlaTyrLysSerArgAlaArgAlaValArgAlaSerLeu	2953
DB	7995	ACACATATGACATACATCATCGCCAGAACAGAGGTTTCTGCTTACGCTGAGAG	8054	DB	9075	CTTGGGCTCCGCGCTTGCAGCTTGCAGACACCGGCGCGGAGCGTCCGCGCTAGCGTT	9134
QY	2594	GlyGlyLysLysAlaAlaArgLeuIleValTyrProAspLeuGlyValArgValCysGlu	2613	QY	2954	IleSerArgGlyGlyArgAlaAlaValCysGlyArgTyrLeuPheThrAlaValLys	2973
DB	8055	GGGGGTCTGTAAGCCAGCTGCTCATCTGCTGTTCCCGACCTGGGCGTGGCGTGGCGAG	8114	DB	9135	CTGTCCAGAGGAGGCGGCTGCTCATATGTGGCAGTACTCTTCAACTGGGCGAGTAA	9194
DB				QY	2974	ThrLysLeuLysLeuThrProLeuProGluAlaArgLeuLeuAspLeuSerSerTrpPhe	2993

Db 9195 ACAAGCTCAAACTCACTCCAAATAGCGCGCTGGCGGCTGGACTTGTCCGGTTGGTTC 9254
 Qy 2994 ThrValGlyAlaGlyGlyGlyAspIleTyrHisSerValSerArgAlaArgProArgLeu 3013
 Db 9255 ACCGCTGGCTACAGCGGGGAGACATTTATCAGCGGTGTCTCATGCCCGCGCCCGCTGG 9314
 Qy 3014 LeuLeuPheGlyLeuLeuLeuPheValGlyValGlyLeuPheLeuLeuProAlaArg 3033
 Db 9315 TTCTGGTTTTCCTACTCTCTGCTCGCTGCAGGGGTAGGCATCTACCTCTCCCAACCGA 9374

RESULT 15

US-09-034-756-1
 ; Sequence 1, Application US/09034756
 ; Patent No. 6392028
 ; GENERAL INFORMATION:
 ; APPLICANT: RICE, CHARLES et al.
 ; TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C
 ; TITLE OF INVENTION: VIRUS (HCV) AND USES THEREOF
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: HOWELL & HAFERKAMP, L.C.
 ; STREET: 7733 FORSYTH BLVD., SUITE 1400
 ; CITY: ST. LOUIS
 ; STATE: MO
 ; COUNTRY: USA
 ; ZIP: 63105
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/034,756
 ; FILING DATE: 04-May-1998
 ; CLASSIFICATION: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: HOLLAND, DONALD R.
 ; REGISTRATION NUMBER: 35,197
 ; REFERENCE/DOCKET NUMBER: 6029-4831
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 314-727-5188
 ; TELEFAX: 314-727-6092
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 9646 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; HYPOTHEICAL: NO
 ; ANTI-SENSE: NO
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1:

Alignment Scores:
 Pred. No.: 0 Length: 9646
 Score: 11973.00 Matches: 2175
 Percent Similarity: 83.29% Conservative: 357
 Best Local Similarity: 71.55% Mismatches: 472
 Query Match: 74.27% Indels: 36
 DB: 4 Gaps: 8

US-09-980-559-2 (1-3033) x US-09-034-756-1 (1-9646)

Qy 1 MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGln 20
 Db 342 ATGAGCAGCAATCTTAAACCTCAAGAAACCAACAGTAACCAACCGTCCGCCACAG 401
 Qy 21 AspValLysPheProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArg 40
 Db 402 GACGTCAAGTTCCGGGTGGCGGTACAGTGTGTGGAGTTTACTTTGTTCGGCGCAGG 461

Qy 41 GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly 60
 Db 462 GGCCTTAGATTGGTGTGGCGGAGAGAGATTCGAGCGGTCCGCAACTCGAGGT 521
 Qy 61 ArgArgGlnProLysAspArgArgSerThrGlyLysSerTrpGlyLysProGly 80
 Db 522 AGACGTGACGCTATCCCAAGGACGTCGCGCCGAGGCGAGGACCTGGGCTCAGCCCGG 581
 Qy 81 TyrProTrpProLeuTyrGlyAsnGlyLeuGlyTyrAlaGlyTyrTrpLeuLeuSerPro 100
 Db 582 TACCCTTGGCCCTCTATGGCAATAGGGTTGGGGTGGGGATGGCTCTCTCTCTCCC 641
 Qy 101 ArgGlySerArgProSerTrpGlyProAsnAspProArgHisArgSerArgAsnValGly 120
 Db 642 CGTGGCTCTCGGCTAGCTGGGCGCCACAGACCCCGCGGTAGGTGCGCAATTTGGGT 701
 Qy 121 LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProValVal 140
 Db 702 AAGGTATCATGATACCTTACGTGCGGCTTCGCGGACCTCATGGGGTACATACGCTGTC 761
 Qy 141 GlyAlaProLeuGlyGlyValAlaArgAlaLeuAlaHisGlyValArgValLeuGluAsp 160
 Db 762 GCGGCGCTCTTGGAGGCGCTGCGAGGCGCTGGCGCATGGGTCTGGGTCTGGAGAC 821
 Qy 161 GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla 180
 Db 822 GCGTGAACATATGCAACAGGAACTTCTCTGGTGTCTCTTCTCTATCTCTCTCTGCGC 881
 Qy 181 LeuLeuSerCysIleThrThrProValSerAlaAlaGluValLysAsnIleSerThrGly 200
 Db 882 CTGCTCTCTTGTGCTGACTGTGCGGCTTCCAGCTACCAAGTGGCAATTTCTCTGGGCTT 941
 Qy 201 TyrMetValThrAsnAspCysThrAsnAspSerIleThrTrpGlnLeuGlnAlaVal 220
 Db 942 TACCATGTACCAATGATTGCGCTTAACTCGAGTATTGTGTAGAGGCGGCGCATCATC 1001
 Qy 221 LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTrpIle 240
 Db 1002 CTGCACACTCGGGGTGTGTCTTCTGGGTTCGCGAGGTTAACTCGAGGTCTGGGTG 1061
 Qy 241 ProValSerProAsnValAlaValGlnArgProGlyAlaLeu---ThrGlnGlyLeuArg 259
 Db 1062 GCGGTGACCCCGCGGTGGCCACC---AGGGACGCGCAAACTCCCCCAACGCGCTTCCA 1118
 Qy 260 ThrHisIleAspMetValValMetSerAlaThrLeuCysSerAlaLeuTyrValGlyAsp 279
 Db 1119 CGTCATATCATCTGCTGTGGGAGCGCCACCTCTGTCTGCGCTCTACGTGGGGGAC 1178
 Qy 280 LeuCysGlyValMetLeuAlaAlaGlnMetPheIleValSerProGlnHisIleTrp 299
 Db 1179 CTGTGGGGTCTGCTTCTTCTTGTGGTCAACTGTTTACCTTCTCTCCAGCGCCACTGG 1238
 Qy 300 PheValGlnAspCysAsnCysSerIleTyrProGlyThrIleThrGlyHisArgMetAla 319
 Db 1239 ACGACGCAAGACTGCAATTTGTTCTATCTATCCCGGCATATAACGGGTCTATCGCATGCA 1298
 Qy 320 TrpAspMetMetMetAsnTrpSerProThrAlaThrMetIleLeuAlaTyrAlaMetArg 339
 Db 1299 TGGGATATGATGATGAACCTGGTCCCTACGCGAGCGTGTGGTGTAGCTCAGTGTCTCCGG 1358
 Qy 340 ValProGluValIleIleAspIleIleSerGlyAlaHisTrpGlyValMetPheGlyLeu 359
 Db 1359 ATCCCAACAGCCATCATGGACATGATCGTGTGTCTACTGGGAGTCTCTGGGGGCATA 1418
 Qy 360 AlaTyrPheSerMetGlnGlyAlaTrpAlaLysValValIleLeuLeuLeuAlaAla 379
 Db 1419 GCGTATTTCTCCATGTGGGGAACCTGGGCAAGGCTCTGGTGTAGTGTCTGTCTATTGTC 1478
 Qy 380 GlyValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeu 399
 Db 1479 GCGTCGACCGGAAACCCACCTCCCGGGGAGTGGCGCGGCGGACCGCGCTGGGCTT 1538
 Qy 400 ThrSerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySer 419

1539	GTGTGTTCTCTTTACACCGCGCCCAAGCAGAACATCCAACTGATCAACCAACCAACGCGAGT	1598
420	TrpHisIleAsnArgThrAlaLeuAenCysAenAspSerLeuHisThrGlyPheIleAla	439
1599	TGGCACATCAATAGCAGCGCCCTTGAACATGCAATGAAGCCTTAAACGGCTGTGGTTAGCA	1658
440	SerLeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCys	459
1659	GGGCTCTTCTATCAGCACAAATTCAACTCTTCAGGCTGTCTCTGAGAGGTGTGCCAGCTGC	1718
460	ArgSerIleGluAlaPheArgValGlyTrpGlyAlaLeuGlnTyrGluAspAsnValThr	479
1719	CGACGCCCTTACCGAATTTTGCCCGCAGGCTGGGGTCTCTATCAGTTATGCGCAACGGA	1772
480	AsnProGluAspMetArgProTyrCysTrpHisTyrProProArgGlnCysGlyValVal	499
1773	ACGGGCTCGACGAACGCCCTTACTGCTGGCACTACCTTCCAAAGACCTTGTGGCAATGTG	1832
500	SerAlaLysThrValCysGlyProValTyrCysPheThrProSerProValValValGly	519
1833	CCGCGAAGAGCGTGTGTGGCCGGTATATGCTTCACTCCAGCCCCGTGTGTGTGGGA	1892
520	ThrThrAspArgLeuGlyAlaProThrTyrThrTrpGlyGluAsnGlnThrAspValPhe	539
1893	ACGACCGACAGGTGCGGGCGCGCTACCTACAGCTGGGTGCAATGATACGGATGCTCTTC	1952
540	LeuLeuAsnSerThrArgProProLeuGlySerTrpPheGlyCysThrTrpMetAsnSer	559
1953	GTTCCTTAAACACCAACGAGGCCACCGTGGGCAATGTGTCGGTTGTACTCGATGAACATCA	2012
560	SerGlyTyrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAla	579
2013	ACTGGATTCCACCAAGTGTGGGAGGCGCCCTTGTGTCTATCGAGGGGTGGGCAAC	2069
580	SerThrAspLeuLeuCysProThrAspCysPheArgLysHisProAspThrThrTyrLeu	599
2070	---AACACCTTGTCTCTGCCCACTGATTGTTTCCGCAAGCATCCGGAAGCCACATACTCT	2126
600	LysCysGlySerGlyProTrpLeuThrProArgCysLeuLeuAspTyrProTyrArgLeu	619
2127	CGGTGCGGCTCCGGTCCCTGGAAATACACCAGGTGCATGGTGCAGTACCCTGATAGGCTT	2186
620	TrpHisTyrProCysThrValAsnTyrThrIlePheLysIleArgMetTyrValGlyGly	639
2187	TGGCACTATCTTGTACCATCAATTACACCATATTCAAGTCAGGATGTACGTGGGAGGG	2246
640	ValGluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGlu	659
2247	GTGAGCACAGGCTGGAAACGCGCTGCAACTGGACGCGGGGCGAAGCGCTGTGATCGGAA	2306
660	AspArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGluTrpAlaIleLeu	679
2307	GACAGGGACAGGTCCGAGCTCAGCCCATTTGCTGTCCACCACACACAGGTGGCAGGTCTT	2366
680	ProCysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisIleuHisGlnAsn	699
2367	CCGTGTCTTTCACGACCCCTGCCAGCTTGTCCACCGGCTCATCCACTCTCACCGAACAC	2426
700	IleValAspValGlnPheMetTyrGlyLeuLeuSerProAlaLeuThrLysTyrIleValArg	719
2427	ATTGTGACGTCGAGTACTTGTAGCGGGTAGGGTCAAGCATCCGCGTCTGGGCCCATTAAG	2486
720	TrpGluTrpValIleLeuLeuPheLeuLeuAlaAspAlaArgValCysAlaCysLeu	739
2487	TGGGAGTAGCTGTCTTCCTGTTCCCTCTGCTTGCAGACGCGCGGCTGTGCTCTGCTTGT	2546
740	TrpMetLeuIleLeuLeuGlyGlnAlaGluAlaAlaLeuGluLysLeuValIleLeuHis	759
2547	TGATGATGTACTATCCCAAGCGAGGGGCTTTGGAGAACCTCTGTATATCTCAAT	2606
760	AlaAlaSerAlaAlaSerCysAsnGlyPheLeuTyrPheValIlePheValAlaAla	779

2607	GCAGCATCCCTGGCGCGGAGCCACAGGCTTGTGTGCTCTCTCTCGTGTTCTTCTTGCTGCTTGGC	2666	
780	TrpTyrIleYsGlyArgValValProLeuAlaThrTyrSerLeuLeuThrGlyLeuThrPsr	799	
2667	TGGTATCTGAAGGGTAGGTGGGTGCCGAGCGGTCTACGCTTCTACGGGATGTGGCT	2726	
800	PheSerLeuLeuLeuLeuAlaLeuProGlnAlaTyrAlaTyrAspAlaSerValHis	819	
2727	CTCTCTCTGCTCTGCTGGCGTTCCTCAGCGGGCAGATACGCACTGGACACGGAGGTGGCC	2786	
820	GlyGlnIleGlyAlaAlaLeuValMetIleThrLeuPheThrLeuThrProGlyTyr	839	
2787	CGTCGTGTGGCGGTTGTCTTGTGGGTATGATGGCGTACTCTCTGCCATATTATC	2846	
840	LysThrLeuLeuSerArgPheLeuTrpTrpLeuCysTyrLeuLeuLeuGlyGluAla	859	
2847	AAGCGCTACATCAGCTGGTGCATGTGGTGGCTTCAGTATTTTCTGACCAGAGTAGAAGCG	2906	
860	MetValGlnGluTrpAlaProProMetGlnValArgGlyGlyArgAspGlyIleIleTrp	879	
2907	CAACTGACGCTGTGGGTTCCTCCCTCAACGTCGCGGGGGCGCGATCGCGTCACTCTTA	2966	
880	AlaValAlaIlePheTyrProGlyValValPheAspIleThrLysTrpLeuLeuAlaVal	899	
2967	CTCATGTGTGTGTACACCCGACTCTGGTATTTGACATCACAAACTACTCTCTGGCCATC	3026	
900	LeuGlyProAlaTyrLeuLeuLysGlyAlaLeuThrArgValProTyrPheValArgAla	919	
3027	TTCGGACCCCTTGGATTCTTCAAGCCAGTTTGCTTAAAGTCCCCCTACTCTGTCGGCGTT	3086	
920	HisAlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyGlyArgTyrValGln	939	
3087	CAAGCGCTTCTCCGATCTCGCGCTAGCGCGGAAGATAGCGGAGGTCAITTCGTGCAA	3146	
940	MetAlaLeuLeuAlaLeuGlyArgTyrThrGlyThrTyrIleTyrAspHisLeuThrPro	959	
3147	ATGGCCATCATCAAGTTTAGGGCGCGCTACTGGCACCTAATGTATATAACATCTCAACCCCT	3206	
960	MetSerAspTrpAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIleIle	979	
3207	CTTCGAGACTGGCGGCACAAAGCGCTGCGAGATCTGGCGTGGCTGTGGACCACTGCTC	3266	
980	PheSerProMetGluLysLysValIleValTrpGlyAlaGluThrAlaAlaCysGlyAsp	999	
3267	TTCTCCCAATGGAGACCAGCTCATACGTGGGGGGCAGATACCGCGCGTGGCGGTGAC	3326	
1000	IleLeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAla	1019	
3327	ATCATCAACGGCTTGGCCGCTCTGTCGCCGTAGGGGCCAGGAGATACTGCTTGGGCCAGCC	3386	
1020	AspGlyTyrThrSerLysGlyTrpSerLeuLeuAlaProIleThrAlaTyrAlaGlnGln	1039	
3387	GACGGAATGGTCTCAAGGGGTGGAGTTGCTTGGCGCCCATCACGGCGTACGCCCAGCAG	3446	
1040	ThrArgGlyLeuLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGluGln	1059	
3447	ACGAGAGCCCTCTAGGGTGTATAATCACACGCTGACTGGCGCGGACAAACCAAGT	3506	
1060	AlaGlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSer	1079	
3507	GAGGGTGGGTCCAGATCGTGTCACTGCTACCCCAACCTTCTCTGGCAACGTGCATCAAT	3566	
1080	GlyValLeuTrpThrValTyrHisGlyAlaGlyAenLysThrLeuAlaGlySerArgGly	1099	
3567	GGGGTATCTGACTGTCTACACGGGGCCGAGCAGGAGCANTCGCATCCCAAGGCT	3626	
1100	ProValThrGlnMetTyrSerSerAlaGlyAspLeuValGlyTrpProSerProPro	1119	
3627	CCTGTCTCATCAGATGATACCAATGTGGACCAAGACCTTGTGGGTGGCCCGCTCTCTCAA	3686	
1120	GlyThrLysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArg	1139	
3687	GGTCTCCCGCTCATTCACACCTTCACCTCGCGCTCTCTCGGACCTTTACTGTGTACGAGG	3746	

1140 AsnAlaAspValIleProAlaArgArgGlyAspLysArgGlyAlaLeuLeuSerPro 1159
 3747 CACGCCGATGTCATCCCGTGGCGCGGAGGTAGTAGCAGGGTAGCTTTCGCC 3806
 1160 ArgProLeuSerThrLeuLysGlySerSerGlyProValLeuCysProArgGlyHis 1179
 3807 CGGCCCATTTCTTACTTGAAGGCTCTCGGGGGTCCGCTGTTGTGCCCGCGGAC 3866
 1180 AlaValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPhe 1199
 3867 GCGTGGGCTATTACGGCGCGGTGTGACCGGTGGAGTGGCTAAGCGGTGGACTT 3926
 1200 IleProValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThr 1219
 3927 ATCCCTGTGGAGAACCTAGAGACAACCATGAGATCCCGGTGTTACGGGACACTCTCT 3986
 1220 ProProAlaValProGlnThrThrGlnValGlyTyrLeuHisAlaProThrGlySerGly 1239
 3987 CCACACAGAGTGGCCAGAGCTTCCAGGTGGCCACCTGCTGCTCCACCGGCGGCT 4046
 1240 LysSerThrLysValProValAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn 1259
 4047 AGAGCACAGGTCCTGGCTGCTAGCAGCCAGCGGTACAGGTGTTGGTGTCTCAAC 4106
 1260 ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerLysAlaHisGlyIleAsn 1279
 4107 CCTCTGTGCTGCAAGCTGGCTTGGTGTCTTACATGTCCAGGCCCATGGGTGAT 4166
 1280 ProAsnIleArgThrGlyValArgThrValThrGlyAlaProIleThrThrSerThr 1299
 4167 CCTAATATCAGGACCGGGGTGAGAACATATACCTGGCAGCCCATCAGTACTCCAC 4226
 1300 TyrGlyLysPheLeuAlaAspGlyGlyCysAlaGlyAlaTyrAspIleIleCys 1319
 4227 TAGGGCAAGTCTCTCCGACCGGGGTGCTCAGGAGGTGCTTATGACATAATTTGT 4286
 1320 AspGluCysHisAlaValAspSerThrIleLeuGlyIleGlyThrValLeuAspGln 1339
 4287 GACGAGTGCACCTCCACGAGTGCACATCCATCTTGGGATCGGCTGCTTGCACAA 4346
 1340 AlaGluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProGlySerVal 1359
 4347 GCAGAGACTCGGGGCGGAGACTGTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4406
 1360 ThrThrProHisProAsnIleGluGluValAlaLeuGlyGlnGluGlyIleProPhe 1379
 4407 ACTGTGTCCTCATCTAACATCGAGGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4466
 1380 TyrGlyArgAlaIleProLeuSerTyrIleLysGlyGlyArgHisLeuIlePheCysHis 1399
 4467 TAGGGCAAGCTATCCCTCGAGGTGATCAAGGGGGGAGACATCTCATCTTCTGCCAC 4526
 1400 SerLysLysCysAspGluLeuAlaAlaLeuArgGlyMetGlyLeuAsnSerVal 1419
 4527 TCAGAGAGAGTGCAGCAGCTCGCGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4586
 1420 AlaTyrTyrArgGlyLeuAspValSerValIleProThrGlnClyAspValValVal 1439
 4587 GCTACTACCGGGTCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4646
 1440 AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn 1459
 4647 TCGACCGATGCTCTCATGCTGCTTACGGGCTTACGGGCTTACGGGCTTACGGGCTTAC 4706
 1460 ValAlaValThrGlnValValAspPheSerLeuAspProThrPheThrIleThrGln 1479
 4707 ACCTGTGTCTCAGACAGTTCGATTTTCAGCTTTCAGCTTTCAGCTTTCAGCTTTCAGCT 4766
 1480 IleValProGlnAspAlaValSerArgSerGlnArgArgGlyArgThrGlyArgGlyArg 1499
 4767 ACCTCTCCCGAGATGCTGCTCTCAGGACTCAACCGCGGCGGAGGAGCTGCGCGGGGAG 4826

1500 LeuGlyIleTyrArgTyrValSerThrGlyGluArgAlaSerGlyMetPheAspSerVal 1519
 4827 CCAGGCATCTACAGATTTGTGGCACCGGGGAGCGCCCTCCGGCATGTTCCGACTCGTCC 4886
 1520 ValLeuCysGluCysTyrAspAlaGlyAlaAlaTyrTyrGluLeuThrProSerGluThr 1539
 4887 GTCTCTGTGAGTGTATGACCGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4946
 1540 ThrValArgLeuArgAlaTyrPheAsnThrProGlyLeuProValCysGlnAspHisLeu 1559
 4947 ACAGTTAGCTACGAGCGGTACATGACACCCCGGGCTTCCGTGTGTCAGGACCATCTT 5006
 1560 GluPheTyrGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln 1579
 5007 GAATTTGGGAGGGCTCTTACGGGCTCTCATATAGATGCCACATCTTATATCCAG 5066
 1580 ThrLysGlnSerGlyGluAsnPheAlaTyrLeuThrAlaTyrGlnAlaThrValCysAla 1599
 5067 ACAGAGCAGAGTGGGAGAACTTCTTACCTGGTAGCGTACCAAGCCACCGTGTGCTG 5126
 1600 ArgAlaLysAlaProProSerTrpAspValMetTyrLysCysLeuThrArgLeuLys 1619
 5127 AGGGCTCAAGCCCTCCCATCTGGGACAGATGTGGAGTGTGTGTGTGTGTGTGTGTGT 5186
 1620 ProThrLeuValGlyProThrProLeuLeuTyrArgLeuGlySerValThrAsnGluVal 1639
 5187 CCCACCTCCATGGGCAACACCCCTGTCTATACAGACTGGGCTGTTCAGAAATGAAGTC 5246
 1640 ThrLeuThrHisProValThrLysTyrIleAlaThrCysMetGlnAlaAspLeuGluVal 1659
 5247 ACCCTGACGACCAATCAACAAATACATCATGACATGTGCGGCCACCTGGAGGTC 5306
 1660 MetThrSerThrTrpValLeuAlaGlyGlyValLeuAlaAlaAlaTyrCysLeu 1679
 5307 GTACAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5366
 1680 AlaThrGlyCysValCysIleIleGlyArgLeuHisIleAsnGlnArgAlaValAla 1699
 5367 TCAACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5426
 1700 ProAspLysGluValLeuTyrGluAlaPheAspGluMetGluGluCysAlaSerArgAla 1719
 5427 CCTGACAGGAGGTTCTTACAGGAGTTCGATGAGATGGAAGAGTGTCTTCAGCACTTA 5486
 1720 AlaLeuIleGluGluGlyGlnArgIleAlaGluMetLeuLysSerLysIleGlnGlyLeu 1739
 5487 CCGTACATCGACAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5546
 1740 LeuGlnGlnAlaSerLysGlnAlaGlnAspIleGlnProThrValGlnAlaSerTrpPro 1759
 5547 CTGACAGCGCTTCCCGCAACGAGGTTATCACCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 5606
 1760 LysValGluGlnPheThrAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu 1779
 5607 AAATCGAGGTTCTTGGGCGGAGCAGCATGTGGAATTTTCATGAGTGGATACAAATCTTG 5666
 1780 AlaGlyLeuSerThrLeuProGlyAsnProAlaValAlaSerMetMetAlaPheSerAla 1799
 5667 GCGGGCTGTCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5726
 1800 AlaLeuThrSerProLeuSerThrThrIleLeuLeuAsnIleLeuGlyGlyTyr 1819
 5727 GCGGTCAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5786
 1820 LeuAlaSerGlnIleAlaProProAlaGlyAlaThrGlyPheValValSerGlyLeuVal 1839
 5787 GTGGCTGCCAGCTGCCCGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5846
 1840 GlyAlaAlaValGlySerIleGlyLeuGlyLysValLeuValAspIleLeuAlaGlyTyr 1859
 5847 GCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5906
 1860 GlyAlaGlyIleSerGlyAlaLeuValAlaPheLysIleMetSerGlyGlyLysProSer 1879

2954	IleSerArgGlyGlyArgAlaIaValCysGlyArgTyrLeuPheAsnTrpAlaValLys	2973
9135	CTGTCCAGAGGAGGAGCGGCTGCCATGTGCGCAAGTACCTCTTCAACTGGCGCAGTAAGA	9194
2974	ThrIysLeuLysLeuThrProLeuProGluAlaArgLeuLeuAspLeuSerSerTrpPhe	2993
9195	ACAAGCTCAAACTCATTCCCAATAGCGCGCGCTGGCGCGCTGGACTTGTGCGGTGGTTC	9254
2994	ThrValGlyAlaGlyGlyYAspIleTyrHisSerValSerArgAlaAargProArgLeu	3013
9255	ACGGCTGGCTACAGCGGGGGAGACATTATTCACAGCGTGTCTCATGCCCGCGCGCGCTGG	9314
3014	LeuLeuPheGlyLeuLeuLeuLeuPheValGlyValGlyLeuPheLeuLeuProAlaArg	3033
9315	TTCTGGTTTGGCTACTCTGTCTGCTCGAGGGGTAGGCATCTACTCTCTCCCAACGGA	9374

Copied from 09540843 on 05/19/2004

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 11, 2004, 23:17:49 ; Search time 11827 Seconds
(without alignments)
6232.813 Million cell updates/sec

Title: US-09-980-559-2
Perfect score: 16120
Sequence: 1 MTNPKPQKTKRNTNRQ.....LLFGLLLFVGVLFLPAR 3033

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2/USFIO.spool/US0980559/runat_09022004.141954.26410/app_query.fasta_1.3207
-DB=EST -Qfmt=fastap -SUFFIX=est -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pc-o -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0980559 @CGN 1.1 9743 @runat_09022004.141954.26410 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -JONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estm.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pin.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	553	3.4	488	9	AV755731
C 2	430.5	2.7	492	9	AV758366
C 3	200.5	1.2	9330	11	AK090118 Mus muscu
C 4	200.5	1.2	9330	11	AK090118 Mus muscu
C 5	185	1.1	2891	11	CNSLTI164
C 6	181	1.1	3865	11	AK032205 Mus muscu
C 7	179.5	1.1	2861	11	AK004685 Mus muscu
C 8	179.5	1.1	5068	11	BC032073 Mus muscu
C 9	175	1.1	2837	11	BC032614 Homo sapi
C 10	173.5	1.1	3929	11	BC049191 Mus muscu
C 11	167	1.0	4257	11	BC032156 Mus muscu
C 12	166.5	1.0	2798	11	AK083804 Mus muscu
C 13	166	1.0	4665	11	AK080374 Mus muscu
C 14	165.5	1.0	3921	11	BC020496 Homo sapi
C 15	165	1.0	4640	11	BC037307 Homo sapi
C 16	162	1.0	4176	11	BC021544 Homo sapi
C 17	162	1.0	4185	11	BC014631 Homo sapi
C 18	161.5	1.0	3965	11	AK089263 Mus muscu
C 19	161	1.0	1255	10	BG029342 Mus muscu
C 20	161	1.0	3291	11	BC046781 Mus muscu
C 21	161	1.0	8315	11	AK090135 Mus muscu
C 22	160.5	1.0	2897	11	BC020381 Homo sapi
C 23	158.5	1.0	2820	11	AK044744 Mus muscu
C 24	158	1.0	2851	10	AY096240 AY096240
C 25	157	1.0	2525	11	AK046543 Mus muscu
C 26	157	1.0	2525	11	AK046552 Mus muscu
C 27	157	1.0	2525	11	AK046553 Mus muscu
C 28	156	1.0	3453	11	BC039131 Homo sapi
C 29	156	1.0	4137	11	BC024101 Homo sapi
C 30	155	1.0	2822	11	AK082152 Mus muscu
C 31	155	1.0	5827	11	BC038226 Homo sapi
C 32	154.5	1.0	1835	29	CC195441 CH261-184
C 33	153.5	1.0	3336	11	AK044366 Mus muscu
C 34	153	0.9	8539	11	AK090137 Mus muscu
C 35	152.5	0.9	3258	11	AK032300 Mus muscu
C 36	152.5	0.9	4106	11	AK036563 Mus muscu
C 37	152	0.9	1087	29	AG082158 Pan trogl
C 38	151.5	0.9	1534	12	B1198766 602759794
C 39	151.5	0.9	1987	10	AW888223 MXRA8 Hum
C 40	151.5	0.9	4350	11	AK044188 Mus muscu
C 41	151	0.9	1142	13	B0722517 AGENCOURT
C 42	150.5	0.9	763	14	CB627321 OSTIEB02D
C 43	150.5	0.9	2000	10	BG033303 60298558
C 44	150.5	0.9	4574	11	AK051511 Mus muscu
C 45	150.5	0.9	5157	11	AK040971 Mus muscu

ALIGNMENTS

RESULT 1
AV755731/c
LOCUS AV755731 488 bp mRNA linear EST 19-OCT-2000
DEFINITION AV755731 BM Homo sapiens cDNA clone EMPAKB03 5', mRNA sequence.
ACCESSION AV755731
VERSION AV755731.1 GI:10913579
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 488)
AUTHORS Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,

Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and Chen, Z.
Homo sapiens cDNA BM clones
Unpublished

TITLE JOURNAL COMMENT

Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES source

Location/Qualifiers
1. .488
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMFAK03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/notes="Vector: pTriplEx2; Site 1: sfiIA; Site 2: sfiIB"
BASE COUNT 116 a 134 c 137 g 97 t 4 others
ORIGIN

Alignment Scores:
Pred. No.: 1.11e-31 Length: 488
Score: 553.00 Matches: 113
Percent Similarity: 74.25% Conservativeness: 11
Best Local Similarity: 67.66% Mismatches: 29
Query Match: 3.43% Indels: 14
DB: 9 Gaps: 2

US-09-980-559-2 (1-3033) x AV755731 (1-488)

Qy 266 ValMetSerAlaThrLeuCyssSerAlaLeuTyrValGlyAspLeuCyGlyValMet 285
Db 472 GTGGTGTACACCTCGCTCTGCTCAGCTCTCTAGCTGTGGGACCTCTCGCAGGAGTGATG 413
Qy 286 LeuAlaAlaGlnMetPheIleValSerProGlnHisHsrPheValGlnAspCysAsn 305
Db 412 CTTGAGCTTTCAGCTG---ATCATCTGGCTTCAGCACCACCATGAGTTTGTGCATGATGCAC 356
Qy 306 CysSerIleTyrProGlyThrIleThrGlyHisArgMet-AlaTrpAspMetMetMetAs 325
Db 355 TGCTCCATCTATCTCGCGCCATCCTGACATGGACCCGATGATGACATGGACATGATGAA 296
Qy 325 nTrpSerProThrAlaThrMetIleLeuAlaTyrAlaMetArgValProGluValIleI 345
Db 295 CTGGCTGTGACCGCTGCTATGATCATGGGCTACGCAATCGCGTTCTCGAGTCATCAT 236
Qy 345 eAspIleIleSerGlyAlaHisTrpGlyValMetPheGlyLeuAlaTyrPheSerMetG 365
Db 235 AGATATCATCAGCGGGCTCCTGCTGGCGCTCATGTTCCGGTTAGCTTCTCTATGCA 176
Qy 365 nGlyAlaTrpAlaLysValValIleLeuLeuLeuAlaAlaGlyValAspAlaArgth 385
Db 175 GGGAGCGTGGCGAAGTCTGCTGTCATCTTCTGTGACCGCTGGGGTACACGGAGCAC 116
Qy 385 rHisThrValGlyGlySerAla-AlaGlnThrThrGlyArgLeuThrSerLeuPheAspM 405
Db 115 CCAGTCCACAGTGGCACTGTGACGGCGCGCCAGCAGTTTTCGTGAGCTTCTGCACCC 56
Qy 405 etGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTrpHisIleAsnArg 425
Db 55 CTGGCGGTAAACAGACATCCACTCAT----- 27
Qy 425 hrAlaLeuAsnCysAsn 430
Db 26 --ANGATCNATTGCAAT 12

RESULT 2

AV758366/C
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AV758366 492 bp mRNA linear EST 19-OCT-2000
AV758366 BM Homo sapiens cDNA clone BMFAK03 5', mRNA sequence.
AV758366
AV758366
EST.
AV758366.1 GI:10916214
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 492)
Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and Chen, Z.

REFERENCE AUTHORS

TITLE JOURNAL COMMENT

Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES source

Location/Qualifiers
1. .492
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BMFAK03"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="BM"
/notes="Vector: pTriplEx2; Site 1: sfiIA; Site 2: sfiIB"
BASE COUNT 124 a 128 c 125 g 112 t 3 others
ORIGIN

Alignment Scores:
Pred. No.: 1.98e-22 Length: 492
Score: 430.50 Matches: 97
Percent Similarity: 70.25% Conservativeness: 14
Best Local Similarity: 61.39% Mismatches: 44
Query Match: 2.67% Indels: 3
DB: 9 Gaps: 1

US-09-980-559-2 (1-3033) x AV758366 (1-492)

Qy 264 MetValMetSerAlaThrLeuCyssSerAlaLeuTyrValGlyAspLeuCyGlyGly 283
Db 479 ATAGTTGTGCATACACACATGCTGTGTCAGCTCAGCTGTGGACCTCTGCGTTGGG 420
Qy 284 ValMetLeuAlaAlaGlnMetPheIleValSerProGlnHisHsrPheValGlnAsp 303
Db 419 GTATCCTTTCAGCCCACTG---ATTATCTTCAGCACAACATTTGGTTTGTGCAAGAA 363
Qy 304 CysAsnCysSerIleTyrProGlyThrIleThrGly-HisArgMetAlaTrpAspMet 323
Db 362 TGCACTGTCTATCTTATCTCTGCTCATCTGCTGACTACAGTATGGCATAGGCTATGAT 303
Qy 323 tMetAsnTrpSerProThrAlaThrMetIleLeuAlaTyrAlaMetArgValProGluVa 343
Db 302 GATGAATCTGTCGCGACCGGTTTCCATGATCTGCGGTACGCAATCGCGTTCTCTGAAGT 243
Qy 343 lIleIleAspIleIleSerGlyAlaHisTrpGlyValMetPheGlyLeuAlaTyrPhe 363
Db 242 CCTCATGATATCATAGCTGGGCACACTGGCGCTCATGTTGGCTCAGTTTCTTTCAC 183
Qy 363 rMetGlnGlyAlaTrpAlaLysValValIleLeuLeuLeuAlaAlaGlyValAspAl 383
Db 182 AATGCAGGAGCGTTTGGCCAAAGTAGCTGTCTCATCCATCTGTTCCACCGCTGGGTTGATGC 123

QY 383 aArgThrHisThrValGlyGlySerAlaAlaGln-ThrThrGlyArgLeuThr-SerLeuP 403
 Db 122 GAGCAACCGATCATCAGTGTGATCATCTGCGCGGACGACAGTTCCTCAGGTGCA 63
 QY 403 heAspMetGlyProArgGlnLysIleGlnLeuValAenThrAenGlySer 419
 Db 62 TAATGCTGCGCTAAGCAGTACATCATCAGTCTATGATGCGTCAATGGCAAT 13

RESULT 3
 AK090118
 LOCUS AK090118.1 GI:26105739
 DEFINITION Mus musculus blood RCB-0035 WEHI-3 CDNA, RIKEN full-length enriched library, clone:G431001B21 product:Notch gene homolog 1, (Drosophila), full insert sequence.

ACCESSION AK090118
 VERSION AK090118.1
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE 1
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636

REFERENCE 2
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159

REFERENCE 3
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--184-format sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861

REFERENCE 4
 AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409, 685-690 (2001)
 REFERENCE 5

REFERENCE 6
 AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 JOURNAL Nature 420, 563-573 (2002)
 REFERENCE 7
 AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, J., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, Y., Tanaka, T., Tomaru, A., Toya, T., Yaeunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE Direct Submission
 JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
 COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site for further details.
 URL:http://genome.gsc.riken.go.jp/
 URL:http://fantom.gsc.riken.go.jp/.

FEATURES
 source
 1. 9330
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="BALB/c"
 /db_xref="PANTOM DB:G431001B21"
 /db_xref="taxon:10090"
 /clone="G431001B21"
 /cell_line="RCB-0035 WEHI-3"
 /tissue_type="blood"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 241..7835
 /note="Notch gene homolog 1, (Drosophila) (MGI:97363, GB|NM_008714, evidence: BLASTN, 98%, match=8058) putative"
 9314..9319
 /note="putative"
 9330
 /note="putative"
 2787 c 2587 g 1949 t

BASE COUNT 2007 a 2787 c 2587 g 1949 t
 ORIGIN

Alignment Scores:
 Pred. No.: 0.00339 Length: 9330
 Score: 200.50 Matches: 518
 Percent Similarity: 28.74% Conservative: 250
 Best Local Similarity: 19.39% Mismatches: 857
 Query Match: 1.24% Indels: 1056
 DB: 11 Gaps: 136

US-09-980-559-2 (1-3033) x AK090118 (1-9330)

QY 183 SerCysIleThrThrProValSerAlaAlaGluValLysAsnIleSerThrGlyTyrMet 202
 Db 1396 AACTGTGACACCAACCTGTCTCAACGGCAAGCAATCTGCACCTGCCCTCGGGTAC--- 1452
 QY 203 ValThrAsnAspCysThrAsnAspSerIleThrTrpGlnLeuGlnAlaValLeuHis 222
 Db 1453 -----ACAGGCGCAGCTGCGAGCAGGACGTGAGTGCTCTG--- 1494
 QY 223 ValProGlyCysValProCysGluLysValGly--AsnAlaSerGlnCysTrpIleProVa 242
 Db 1495 -----GGTGCCACCTGTGAGCAGCAGCGCAATGCTCAACACATGTTCTTTT 1548
 QY 242 lSerProAsnVal-----AlaValGlnArgProGlyAla-----LeuThrGlnG 257
 Db 1549 GAGTGCCAGTGTCTCAGGGCTACACGGGACCCCGCTGTGAGATTGATGTTAATGAGTGC 1608
 QY 257 yLeuArgThrHisIleAspMetValValMet-----SerAl 269
 Db 1609 ATCTCCACCCATGTGAGAATGATGCACTTGCCTGGACAGATGGGGAGTTCCAATGC 1668
 QY 269 aThrLeuCysSerAlaLeuTyraValGlyAspLeuCysGlyGlyValMet----- 285
 Db 1669 ATATGTGTGCCAGGTATGAAGGTGTATACTGTGAATCAACACCGATGAGTGCAGCAGC 1728
 QY 285 ----- 285

Db	1729	AGCCCTGCTCTGCACAAATGCGCACTGCATGCAACAGATCAATGAGTGTCCAAATGTCAGTGC	1788
Qy	286	-----LeuAlaAlaGlnMetPheIleValSerProGlnHisI	298
Db	1789	CCGAGGCTTCAACGGGCACCTGTGCGCATATGATGTGATGATGTGCGCACACACCA	1848
Qy	298	8-----TrpPhe-----ValGlnAspCysAsnCysSe	307
Db	1849	TGCAAGAACGGTGCACAGTGCCTGGATGGGCCCAACCTATATCTCGGTGTGTACAGAA	1908
Qy	307	rIleTyrProGlyThrIleThrGlyHisArgMetAlaTrpAspMetMetAenTrpSe	327
Db	1909	GTTTACACAGGA-----	1921
Qy	327	rProThrAlaThrMetIleLeuAlaTyrAlaMetArgValProGluValIleIleAspII	347
Db	1922	CCCACTGCCAAGTGGACATTCAGCAGTGTGACCCCTGACCCCTGCCACTATGTTCTGT	1980
Qy	347	eIleSerGlyAlaHisTrpGlyValMetPheGlyLeuAlaTyrPheSerMetGlnGlyAl	367
Db	1981	A-----AGGATGGT	1989
Qy	367	aTrpAlaLysValValIleLeuLeuLeuAlaGlyValAspAlaArgThrHisTh	387
Db	1990	GTGGCCACCT-----TTACCTGCCTGTGCCAGCCGCTACACAGC	2031
Qy	387	rVal-----GlyGlySerAlaAlaGlnThrThrGlyArgLeuThrSerLeuPheAs	404
Db	2032	CATCATCTGTGAGACCAACATCATGATGAGTGCACCAACCGCCGCTGCCGATGGGGCACC	2091
Qy	404	pMetGlyProArgGlnIleGlnLeuValAsnThrAsnGlySer-----	419
Db	2092	TGCAGGACCGTGACAACTCTACCTCTGCTTATGCTTATGCTCAAGGGAACCAACAGGGGCCCACT	2151
Qy	420	-----TrpHisIleAsnArgTh	425
Db	2152	GTGAGATCAACCTGGATGACTGCGCCAGCAACCCCTGTGACTCTGGCACCTGTCTGG-AC	2210
Qy	425	rAlaLeuAsnCys---AsnAspSerLeuHisThrGlyPheIleAlaSerLeuPheThr	444
Db	2211	AAGATTGATGCTACGAATGT-GCCTGTGAACCAAGGCTACACAGGAAGCATGTGTAACT	2269
Qy	444	rHis-----SerPheAsnSerSerGlyCysProGluArgMe	456
Db	2270	CAACATTGACAAATGTGGGSCACCCCTGCCCAACAGGGGCACTGTGAGGATGGCAT	2329
Qy	456	tSer-----AlaCysArgSerIleGluAlaPheArgValGlyTrpGlyAlaLeuGlnTy	474
Db	2330	CGCGGGCTTCACTTGGCGCTGCCCGAGGGCTACCATGACCCCACTGCTCCGAGGT	2389
Qy	474	rGluAspAsnValThrAsnProGluAspMetArgProTyrCysTrpHisTyrProProAr	494
Db	2390	CAACGAGTGCAACAGTAACCC-----TGCAATCCACGAGGCTTGGCCG	2431
Qy	494	gGlnCysGlyValValSerAlaLysThrValCysGlyProValTyrCysPheThrProSe	514
Db	2432	GGAT---GGCCTCAATGGTCAAGTGTGACTGTGCCCT-----	2468
Qy	514	rProValValValGlyThrThrAspArgLeuGlyAlaProThrTyrThrTrpGlyGluAs	534
Db	2469	-----GGGTGGAGTGGAAAC	2482
Qy	534	nGluThrAspValPheLeuLeuAsnSerThrArgProProLeuGlySerTrpPheGlyCy	554
Db	2483	AAACTGTGACATCAACACACAGAGTGTGAGTGTCCAAACCTTGTGTCAACGGTGGCACCTG	2542
Qy	554	sThrTrpMetAsnSerGlyTyrThrLysThrCys-----GlyAlaProPr	570
Db	2543	CAAGGACATG---ACCACTGCTACGATGTGACCTGTGGAGAGGCTTCAGTGGCCCTAA	2599
Qy	570	oCysArgThrArgAlaAspPheAsnAlaSerThrAspLeuLeu-----	584
Db	2600	TTGCCAGACCAACATCAACGAATGTGCTCTCCAAACCCCTGCAACAGGGGACCTGCAAT	2659
Qy	585	-----CysProThrAspCysPheArgLysHisProAs	595
Db	2660	TGATGATGTCGTGATACAGTGAACCTGTCTCTGCCATATACAGGAGCCACGTGTGA	2719
Qy	595	pThrThrTyrLeuLysCysGlySerGlyPro-----	605
Db	2720	GGTGGTGTGGCCCCCATGTGTACCAAGCCCTGCAAAACAGCGGGGTATGCAAGGAGTC	2779
Qy	606	-----TrpLeuThrProArgCysLe	612
Db	2780	TGAAGACTATGAGAGTTTTTCTGTCTGTCTGCCACAGCTGCGCAAGGTCAAAACCTGCGA	2839
Qy	612	uIleAspTyrProTyrArgLeuTrpHisTyrProCysThrValAsnTyrThrIlePheLy	632
Db	2840	GGTTGACATC---AATGAGTGTGTGAAGCCCATGT-----	2873
Qy	632	sIleArgMetTyrValGlyGlyValGluHisArgLeuThrAlaAlaCysAsnPheThrAr	652
Db	2874	-----CGCCATGGGGCTCTCTGCCAGAACACCAA	2902
Qy	652	gGlyAsp---ArgCysAsnLeuGluAspArg---AspArgSerGlnLeu-----	666
Db	2903	TGGCAGTACCGCTGCT-CTGCCAGCGCGCTATACAGGTGCGCACTGTGAGAGTGACA	2961
Qy	667	-----SerProLeuLeuHisSerThrThrGluTrpAlaIleLeu-----	679
Db	2962	TGATGACTGCGCGCCCAACCGTG-----TCACAATGGGGTTCCTGACCGC	3009
Qy	680	-----ProCysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisLeuHi	697
Db	3010	ATGGCATCAACACAGCTTCTGCGACTGCTGCCGGCTTCAGGGTGCCTTCT-GTGAG	3068
Qy	697	gSlnAsnIleValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrLysTyrIl	717
Db	3069	GAGGACATCAATGAATGTG-----CCAGCAATCCCTGCCAAAAT	3107
Qy	717	eValArg-----Trp-----GluTrpValIleLe	725
Db	3108	GGTGCAATTGCACTGCTGTGGACAGCTACACATGACCTGCCCCCTGGGCTTCAAT	3167
Qy	725	u-----LeuPheLeuLeuAlaAspAlaArgValCysAl	737
Db	3168	GGCATCCACTGCGAGAACACACACCTGACTGTACTGAGAGCTCCTGCTTCAATGGTGT	3227
Qy	737	aCysLeuTrpMetLeuIleLeuLeuGlyGlnAlaGluAlaLeuGluLysLeuValIl	757
Db	3228	ACCTGTGTGGATGGTA-----TCAACTCCTTTCACCTGTCTG	3263
Qy	757	eLeuHisAlaAlaSerAlaAlaSerCysAsnGlyPheLeuTyrPheValIle-Phe---	775
Db	3264	TGTCCACCTGGCTTCAAGGAGAGCTACTGTGAGTATGATGTCAATGAGTGTGATTACCGG	3323
Qy	776	--PheValAlaAlaTrpTyrIleLysGlyArgValValProLeuAlaThrTyrSerLeuT	795
Db	3324	CCCTGTCTGCAGCGTGTACTCTGCCAAGACAGCTGTTACTTATAAGTGTACTTGCCCA	3383
Qy	795	hrGlyLeu---TrpSerPheSerLeuLeuLeuAlaLeuProGlnGlnAlaTyrAlaT	814
Db	3384	CAGGGCTACACTGGTCT-----CAACTGCCAGAACCT-----	3415
Qy	814	yrAspAlaSerValHisGlyGlnIleGlyAlaAlaLeuLeuValMetIleThrLeuPheT	834
Db	3415	-----	3415
Qy	834	hrLeuThrProGlyTyrLysThrLeuLeuSerArgPheLeuTrpTrpLeuCysTyrLeuL	854
Db	3416	-----TGTGCGCTGG	3425
Qy	854	euThrLeuGlyGluAlaMetValGlnGluTrp-----Alap	866
Db	3426	TGGCACTGCGCTCC-----CTGCAAGAATGGTGGCAGGTGTGGCAGACCAACACCGCAG	3479

QY 866 roProMet-GlnValArgGlyGlyArgAspGly-----IleIle 878
 DB 3480 TACCACCTGTGAGTGGCGGAGCGGCTGCACTGGCGTCAACTGGCGACGTGCTCAGTGTGCTCC 3539
 QY 879 TrpAlaValAlaIlePheTyrProGlyValValPheAspIleThrIleThrIleLeuAla 998
 DB 3540 TGTGAGTGGCTGCACAGAGCGAGGCAAT-----GACGTCACT-----CTCCTGTGC 3587
 QY 899 ValLeuGlyProAlaTyrLeuLeuLysGlyAlaLeuThrArgValProTyrPheValArg 918
 DB 3588 CAGCATGGAGGGCTCTGTGTGGATAGGGA----- 3617
 QY 919 AlaHisAlaLeuLeuArgMetCysThrMetAlaArgHis-LeuAlaGlyGlyArgTyrVa 938
 DB 3618 -----GATAAACAATTACTGCCACTGCCAGCGAGCG 3647
 QY 938 lGlnMetAlaLeuLeuAlaLeuGlyArgTyrThrGlyThrIleTyrAspHisLeuTh 958
 DB 3648 TACACGGGAGCTACTGTGAGGACGAGGTGGACGAGTGTCT-----CACCTAAC 3695
 QY 958 rProMetSerAspTipAlaAlaSerGlyLeuArgAspLeuAlaVal---AlaValGluPr 977
 DB 3696 CCCT-----GCCAGATGGAGCTACTCTGCACGACTATCTCGGCGGCTTTTCC 3743
 QY 977 oIleIlePheSerProMetGluLysLysValIleValTrpGlyAla----- 992
 DB 3744 T-----GCAAGTGTGTGGCTGGCTTACCATTGGGTCTAAC 3776
 QY 993 -----GluThrAlaAlaCysGlyAspIleLeuHisGlyLeuProValSerAl 1008
 DB 3777 TGTCTCCGAGGAGATCAACAGTGCCTGT-----CCCAGCCCTGC 3815
 QY 1008 aArgLeuGlyArgGluValLeuLeu-----GlyProAlaaspGly-- 1021
 DB 3816 CAGATGGGGGTACCTGCATGTATCTGACCAACTCTACAGTTCTTCCGCCCGGGGG 3875
 QY 1022 -----TyrThrSerLysGly----- 1026
 DB 3876 ACACAGGGTGTACACTGTGAGATCAATGTTGATGACTGCCATCCCGCTTGACCCCTGCC 3935
 QY 1027 -----TrpSerLeuLeuAlaProIleTh 1034
 DB 3936 TCCCGAAGCCCAAGTGTCTTCAACAATGGCACCTGTGTGGACCGAGTGGGTGCTATACC 3995
 QY 1034 rAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyThrIleValValSerMetThrGlyAr 1054
 DB 3996 TGCACCTGCCCCACCGAGTGTGCGGGAGCGGTGTGAGGGTGTATGTCAATG----- 4047
 QY 1054 gAspLysThrGluGlnAlaGlyGluIleGlnValLeuSerThrValThrGlnSerPheLe 1074
 DB 4048 -----AATGTCCTCCACCCCTGTGACCCACGTGGC 4079
 QY 1074 uGlyThrSerIleSerGlyValLeuThrThrValTyrHisGlyAlaGlyAsnLysThrLe 1094
 DB 4080 ACCCAGACCTGTGTGCGAGCGTGTAAATGACTTCCACTCGGAGTGGCGGGTGGCCACACT 4139
 QY 1094 uAlaGlySerArgGlyProValThrGlnMetTyrSerSerAlaGluGlyAspLeu----- 1112
 DB 4140 GGACGCCGT-----GTGAGTCAGTCATCAATGGCTGCGAGGCGCAACCTTGCMAAG 4190
 QY 1113 -----ValGlyTyrProSerProGlyThrLysSer----- 1123
 DB 4191 AATGGGGGTGTCTGTGCGGTGGCTTCCACACCGCCGGTGGATTCTGTAGTGGCCCT 4250
 QY 1124 -----LeuGluProCysThrCysGlyAlaValAspLe 1134
 DB 4251 GCGGGCTTCGAGGGTCCGCATGTGAGAAATGATGCGGCACCTTGTGGCAGCTTACGCTGC 4310
 QY 1134 uTyrLeuValThrArg-----AsnAlaaspValIleProAlaArgArgGlyAspLys 1152
 DB 4311 CTCACGGTGGTACATGCATCTCGGGCCCGAGTGTAGTCCACCTGCCTATGCTGGGATCC 4370

QY 1152 sArgGlyAlaLeuLeuSer----- 1158
 DB 4371 TTCCACCGGCTGTGAGTGGCAGTTCCAGCCAGCGGCCCTGTGTGGTAGCAACCCCTGC 4430
 QY 1159 -----ProArgProLeuSerThr-----Le 1165
 DB 4431 TACAATCAGGGCACCTGTGAGCCACATCCGAGAACCCCTTTCTACCGCTGTCTATGCCCT 4490
 QY 1165 uLysGlySerSerGly----- 1170
 DB 4491 GCCAAATTCACGGGCTACTGTGCCCATCTCTGGACTACAGCTTCCAGGTGGCGTGGG 4550
 QY 1171 -----GlyProValLeuCysProArgGlyHisAlaVa 1181
 DB 4551 CGCGCATTTCCCCACCGCAGATTGAGGAGGCTGTGAGCTGCCTGAGTGCAGGTGGAT 4610
 QY 1181 lGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIlePr 1201
 DB 4611 GCAGCAATAAGGTCTGCAACCTGCAGT-----GTAATAAT 4646
 QY 1201 oValGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrProPr 1221
 DB 4647 CACGATGTGGTGGATGGTGGCGACTGCTCC-----TCACCTTCAATGACCCCTGG 4700
 QY 1221 oAlaValProGlnThr---TyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyLy 1240
 DB 4701 AAGAACTGCACGCGCTCTCTACAGTGTGGAAGTATTTAGCGCGCCACTGTGACAGC 4760
 QY 1240 sSerThrLysValProValAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnPr 1260
 DB 4761 CAGTGCACCTCGGCCG----- 4776
 QY 1260 oSerValAlaAlaThrLeu---GlyPheGlyAlaTyrLeuSerLysAlaHisGlyIleAan 1279
 DB 4777 -----GCTGCTCTTTGATGGCTTCCAGTCCAGCTCAGCGAGGACAG---TGCAAC 4826
 QY 1280 ProAsnIleArgThrGlyValArgThrValThrGlyAlaProIleThrTyrSerThr 1299
 DB 4827 CCC-----CTGTATGACCCAG 4841
 QY 1300 TyrGlyLys---PheLeuAlaAspGlyGlyCysAlaGlyGlyAla----- 1313
 DB 4842 TACTCAGAGGACCACTTCACTGATGGCCACTCGACCGAGGCTGTAAACAGTCCGCAATGT 4901
 QY 1314 ---TyrAspIleIleCysAspGluCysHisAlaValAspSerThrThrIleLeuGly 1332
 DB 4902 GAGTGGATGGCTAGACTGTCTCAG---CATGTACCGAGCGG-----CTGGCA 4949
 QY 1333 lLeGlyThrValLeuAspGlnAlaGluThrAlaGlyValArgLeuThrValLeuAlaThr 1352
 DB 4950 GCGGCAACCTG-----GTGCTGGTGGTG 4973
 QY 1353 AlaThrProGlySerValThrThrPro----- 1362
 DB 4974 CTGCTTCCACCGCAGCAGCTACGCAACAACCTCTTCCACTTCTGCGGAGCTCAGCCAC 5033
 QY 1363 -----HisProAsnIle---GluGluValAlaLeuGlyGlnGluGlyGluIlePro 1378
 DB 5034 GTGCTGCACCAACCACTGTCTTCAAGCGTGTATGCGAAGGCGCAGCAGATGATCTTCCCG 5093
 QY 1379 PheTyrGly-----ArgAlaIlePro----- 1385
 DB 5094 TACTATGCCCAGGAGAGCTGCGCAAGCACCCCAATCAACGGCTCTACAGTGGGTGG 5153
 QY 1386 -----LeuSerTyrIleLysGlyIleArgHis----- 1394
 DB 5154 GCCACCTCTTCACTGCTTCTGTGTACAGTGTGGCGCGCCAGCGAGGTGAGCCCC 5213
 QY 1395 -----LeuIlePhe-----CysHisSer 1400
 DB 5214 ATGGACATCGTGGCTTCCATTGTCTACTCTGGAGATCGAACCCGGCAATGTGTGAGTCA 5273
 QY 1401 LysLysLysCysAspGluLeuAla-----AlaAlaLeuArgGly 1413

Copied from 09540843 on 05/19/2004

```

QY 2086 ----- 2086
Db 6915 GGTGTGCTGACACTCACCCTGGGATCAGCCACTTGAATGTGGCAGCCAGCTGAGATG 6974
QY 2087 ----- 2088
Db 6975 GCAGCACTGGCTGGAGGTAGCCGGTGGCCCTTTGAGCCACCCCGCCAGCTCTCCAC 7034
QY 2089 ----- 2105
Db 7035 CTGCTGTAGCTTCCAGTGCAGACAGTCTGAGTACCAATGCAGCGGGGTATGAAT 7094
QY 2105 yrlleThrGlyLeuThrAspAsnLeuLysValPro-----CysGlnLeuProSerP 2123
Db 7095 TTCACCGTGG---TGCACCGCAAGCTTGAATGGCCAGTGTGAGTGGCTTCCCGGCTC 7151
QY 2123 roGluPhePhe----- 2126
Db 7152 CAGATGGCATGGTGGCCAGCCAGTCAACCCACTAGCGCGGGGTGTGACGCGGGGCACA 7211
QY 2127 ----- 2142
Db 7212 CTGAGCACAGGAGCTGGCTCCAGCA-----TAGCATGATGGGGCCACTTACACAGC 7265
QY 2142 ysProPhePheArgAspGluValSerPheCysValGlyLeuAsnSerPheValValGly 2162
Db 7266 AGCCT-----CTCCACCAATACCTTGTCCCG 7292
QY 2162 erGlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThrAsp 2182
Db 7293 ATTATTACCA-----GGGC 7307
QY 2182 roSerHisIleThrAlaGluThrAlaAlaArgLeuAlaArgGlySerProProSerG 2202
Db 7308 CTGCCCCAACACAGCTGGCAACACAGCC-----TCACCTGGTGGCAG 7349
QY 2202 luAlaSerSerAlaSerGlnSerAlaProSerLeuArgAlaThrCysThrTh 2222
Db 7350 ACCGAGGAGTGAGCCAGCACTTACACTCCAGC-CTCAGA--ACCTGAGCCACC 7405
QY 2222 isGlyLysAlaTyAspValAspMetValAspAlaLen-----LeuPheW 2237
Db 7406 ATCACAGCCACACTCAGTGTG-AGCTCGGCAGCAATGGGCACCTCGGGCGGAGCTTCT 7464
QY 2237 etGlyGlyAspValThrArgIleGlu-----SerGlySerLysValValValLeu 2254
Db 7465 TGAGTGGGGAGCCAGTCAGGCAGATGTACAACCGCTGGGCCCCAGCAGTCTCCCTGTGC 7524
QY 2254 spSerLeuAspProMetValGluArgSerAspLeuGluProSerIleProSerGlu 2274
Db 7525 ACACATCTCGCC-----CAGAAAGCCAGGCGCTGCCCATCTACTGCCATCTCCA 7578
QY 2274 yrMetLeuPro-----LysLysArgPhePro-ProAlaLeuProAlaTrpAlaArgPro 2291
Db 7579 TGGTCCCAACCCATGACCACTACCACTTCTGAGCCCTCTCCAGCAGATTAATCTCT 7638
QY 2292 AspTyAsnProProLeuValGluSerTrpLysArgProAspTyrGlnProAlaThrVal 2311
Db 7639 -----CTCTCCCTG-----TGACAAACACCC-----CCAGCCAGCAGC 7671
QY 2312 AlaGlyCysAlaLeuProPro----- 2318
Db 7672 TGCAGGTGCCAGAGACCCCTTCTCTACCCCATCTCCCTGAGTCCCTGACCATGGTCCA 7731
QY 2319 -----ProArgLysThrProThr-----ProPro 2326
Db 7732 GCTCTCCCGCATTCACATCTCTGATGTGGTCCGAGGGCATCTCCAGCCCGCCACCA 7791
QY 2327 ProArgArgArgThr 2332
Db 7792 CCATGCGTCCAGATCA 7809

```

RESULT 4
 AK090118/c
 LOCUS
 DEFINITION
 Mus musculus blood RCB-0035 WEHI-3 cDNA, RIKEN full-length enriched library, clone:G431001B21 product:Notch gene homolog 1, (Drosophila), full insert sequence.
 AK090118
 AK090118.1 GI:26105739
 HTC; CAP trapper.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 99279253
 10349636
 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 20499374
 11042159
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishimi, T., Hara, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Hashiwa, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, S., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861
 4
 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)
 5
 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6
 (bases 1 to 9330)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takada, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
 Direct Submission
 Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gscc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
 cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

FEATURES
source

[illegible]

437	QY	he	le	Ala	Ser	Leu	Phe	Tyr	Thr	His	Ser	Phe	Asn	Ser	Ser	Gly	Cys	Pro	Glu	Arg	Met	S	457
6240	Db	GC	AT	CCCT	CC	CA	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	6181
457	QY	ex	Ala	Cys	Arg	Ser	Ile	Glu	Ala	Phe	Arg	Val	Gly	Tyr	Gly	Ala	---	---	---	---	---	---	472
6180	Db	GG	CA	TCC	AG	AT	CT	GT	GG	CC	CT	TT	CC	GA	CG	AG	GA	TCT	GG	AA	GA	CA	6121
473	QY	Gln	Tyr	Glu	Asp	Asn	Val	Thr	Asn	Pro	Glu	Asp	Met	Arg	Pro	Tyr	Cys	Tyr	---	---	---	---	490
6120	Db	CA	AA	AA	CT	CT	GC	AT	GT	AA	CG	GAG	TAC	GG	CC	CA	TG	TT	GT	CT	CT	CT	6061
491	QY	Tyr	Pro	Pro	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	502
6060	Db	TG	GC	CT	CC	CA	AG	CG	CT	TT	GC	AG	CA	CT	GA	AG	AG	TAT	CG	GG	CA	CA	6001
503	QY	Thr	Val	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	515
6000	Db	CG	GT	CT	CC	CG	GT	CG	GT	CT	GT	CT	GT	CT	GT	CT	GT	CT	GT	CT	GT	CT	5941
516	QY	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	529
5940	Db	AG	AT	GAC	AG	CAG	GT	GC	AT	CT	TT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	5881
529	QY	r	Thr	Tyr	Gly	Glu	Asn	Glu	---	---	---	---	---	---	---	---	---	---	---	---	---	---	540
5880	Db	TG	CA	GG	AG	CA	AT	CA	TAG	GG	GT	GT	GA	CC	AT	CT	GG	T	CCT	CG	AA	CA	5821
540	QY	u	Leu	Asn	Ser	Thr	Arg	Pro	Pro	---	---	---	---	---	---	---	---	---	---	---	---	---	557
5820	Db	AG	T	CAG	CA	T	CC	CC	CT	CG	GG	GT	GT	GG	GG	CA	TG	CG	CA	CA	TG	CG	5762
557	QY	ca	sn	Ser	Ser	Gly	Tyr	Thr	Leu	Ser	Cys	Gly	Ala	Pro	Pro	Cys	Arg	Thr	Arg	Ala	Asp	Ph	577
5761	Db	GC	AG	CG	TCC	AG	GT	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	5714
577	QY	ea	sn	Ala	ser	Thr	Asp	Leu	Cys	Pro	---	---	---	---	---	---	---	---	---	---	---	---	586
5713	Db	CT	CAG	GT	CAG	GA	CA	CT	ACT	GG	CT	CT	CA	AA	CC	GA	CT	CT	TG	GT	CT	CT	5654
587	QY	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	602
5653	Db	TCT	CC	CA	CT	CG	TT	CT	GA	TGT	---	---	---	---	---	---	---	---	---	---	---	---	5603
602	QY	ys	er	Gly	Pro	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	618
5602	Db	TT	CAG	GG	CT	TT	CAG	CG	CA	CT	GA	GT	CT	CG	CG	AG	GG	GT	CT	CT	CT	CT	5543
618	QY	rg	Leu	Tyr	His	Tyr	Pro	Cys	Thr	Val	Asn	Tyr	Thr	Ile	Phe	Leu	Val	Arg	Met	Tyr	Val	IG	638
5542	Db	GC	CT	CT	GA	CA	CT	TT	GA	AC	CC	TC	AG	GA	AC	CA	GA	GT	GC	CA	TG	CT	5483
638	QY	ly	Gly	Val	Alu	His	Arg	Leu	Thr	Ala	Ala	Cys	Asn	Phe	Thr	Arg	Gly	Asp	Arg	Cys	Asn	AL	658
5482	Db	CG	GA	CAG	C																		

[illegible]

Db	4320	CAAGTTGAG-----		GCAGGTAAAGCTGCCACAAAGTCCGGGCAT	4282	Db	3586	CACAGGAGAGTGCACGTC-----	AATGCCCTCGCTTCTGTGCAGCCACCTCCACAGGAC	3536
Qy	1043	uLeuGlyThrIleValValSerMetThrGlyArgAspLysThrGluGlnAlaGlyGluIle			1063	Qy	1392	lyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAlaLeuA		1412
Db	4281	CATTCTCACA-----		TGCGGCACCCCT	4261	Db	3535	ACACTGAGCAGCTCCGAGTTCGACGCTCCAGCTCCAGCTCCGACACTCAGAGTGGTACTGC		3476
Qy	1063	eGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGlyValLeuTr			1083	Qy	1412	rgGlyMetGlyLeuAsnSerValAlaTyTyArgGlyLeuAspValSerValIleProTr		1432
Db	4260	CGAAGCCCGCAGGCGCCTCAGATGATCCAGCGGC-----		GGTGTGGAGGC-----	4212	Db	3475	GT---GTTGCTCGCCAGCAGCTCCGACCATCTTCTTGCAC---GGGAGCCGAGTCCGACACAG		3422
Qy	1083	pThrValTyHisGlyAlaGlyAsnLysThrLeuAla-GlySerArgGlyProValThrG			1103	Qy	1432	hrGlnGlyAspValValValAlaThrAspAlaLeuMetThrGlyTyThrGlyAspP		1452
Db	4211	-----CACGGCACACACCCCATCTTGCAGGTTTG-----		CCCTGCAGC	4168	Db	3421	CGCACAAAGGTTCTCGAGTTCAGACGAGTGTAGCCCTGTGGGCAAGT-----		3375
Qy	1103	lnMetTySerSerAlaGlyAspLeuValGlyTrpProSerProProGlyThr---		L	1122	Qy	1452	heAspSerValIleAspCysAsnValAlaValThrGlnValValAspPheSerLeuAspP		1472
Db	4167	CATTGATGACTACTACAGCGGCGTCCAGTG---TGCGCAGCCCGGCACTCGCAGTGA		4111	Db	3374	-----ACCTTATAAGTACCATAGCTGTCTTGGCAGGTACCCGCTGCAGACAGGCG		3323	
Qy	1122	ysSerLeuGluProCysThrCysGlyAlaValAspLeuTyLeuValThrArgAsnAlaA		1142	Qy	1472	roThrPheThrIleThrThrGlnIleValProGlnAspAlaValSerArgSerGlnArgA		1492	
Db	4110	AGTCATTAAACACGCTGCACA-----		CAGT	4087	Db	3322	CGTGAATCACACTCATTGACATCATCTG-----ACAGTAGCTGCCCTGAAGCCAGGT		3269
Qy	1142	spValIleProAlaArgArgArgGlyAspLysArgGlyAlaLeuLeuSerProArgProL		1162	Qy	1492	rgGlyArgThrGlyArgGlyArgLeuGlyIleTyArgTyValSerThrGlyGluArgA		1512	
Db	4086	TCTGGTGCCACGCTGGGTCCAGGGTGGGAGACATTCATTGACATCACCTCCACACC		4027	Db	3268	GCACACAGACAGGTGAAGCA-----GTTGATACCATCCACACAGGTACCACCA		3221	
Qy	1162	euserThrLeuLysGlySerSerGlyGlyProValLeuCysProArgGlyHisAlaValG		1182	Qy	1512	laSerGlyMetPheAspSerValValLeuCysGluCysTySerAspAlaGlyAlaAla-Tp		1531	
Db	4026	GCTCCCGCAGCAAG-----		CCTGGTGGG-----	4004	Db	3220	TTGAAGCAGGAGCTCTCAGTACAGTCAGGTGT-----GTTGTTCTCGCAGTGG		3173
Qy	1182	lyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIleProV		1202	Qy	1531	TyrGluLeuThrProSerGluThrValArgLeuArgAlaTyPheAsnThrProGlyI		1551	
Db	4004	-----		4004	Db	3172	ATGCCATTGAAGCCC-----ACGGGCAGGTACATGTGTAGTGTCCACACAGTCA		3122	
Qy	1202	alGluThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrProProA		1222	Qy	1551	LeuProValCysGlnAspHisLeuGluPheTrpGluAlaValPheThrGlyLeuThrHis		1571	
Db	4003	-----CAGTCCAGGTATAGCCACCACCTCGTCCACA-----		3971	Db	3121	GTGCAATTGGCA-----CCATTTTGGCAGGATTCGTG-----		3089	
Qy	1222	laValProGlnThrTyGlnValGlyTyLeuHisAlaProThrGlySerGlyLysSerT		1242	Qy	1571	IleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnPheAlaTyLeuThr		1591	
Db	3970	-----CAGTGCCATTGTTGAA-GCACTTGGGGCTTCGGGAGGCGGAGG		3929	Db	3082	-----GCACATTCATTGATGTCTCTCCACAG-----		3062	
Qy	1242	hrLysValProValAlaTyAlaGlnGlyTyLysValLeuValLeuAsnProSerV		1262	Qy	1591	AlaTyGlnAlaThrValCysAlaArgAlaAlaProProProSerTrpAspValMet		1611	
Db	3928	TCAAGGGGGGATGGCAGTCAACATTCATCAGTGATACACCTGTGTCCCCCGG		3869	Db	3061	-----AAGGCACCC-----		3053	
Qy	1262	alAlaAlaThrLeuGlyPheGlyAlaTyLeuSerLysAlaHisGlyIleAsnProAsnI		1282	Qy	1611	TrpLysCysLeuThrArg-----		1626	
Db	3868	GGGACGGACACTT-----		3855	Db	3052	TGGAAGCCGGCAGGAGGAGTCCGAGAAGGCTGTGTGATGCCATCGGTGCA-GGA---ACC		2997	
Qy	1282	leArgThrGlyValArgThrValThrThrGlyAlaProIleThrTySerThrTyGlyL		1302	Qy	1626	ProLeuLeuTyArgLeuGlySerValThrAsnGluValThrLeuThrHisProValThr		1646	
Db	3854	-----GTAGGAGTTGGTCAGATCAATGCAAGTACCCCAAT---CTGGCAGGCTGGGAC		3803	Db	2997	CCATTGTACACGGGTGGGGCGGCGAGTCATCGATGTCACTCTCACA-----		2949	
Qy	1302	ysPheLeuAlaAspGlyCysAlaGlyAlaTyAspIleIleIleCysAspGluC		1322	Qy	1646	LysTyIleAlaThrCysMetGlnAlaAspLeuGlu---ValMetThrSerThrTrpVal		1665	
Db	3802	AGGCACCTCGTTGAT-----CTCCTCGGAGCAGTTAGACCCATGGTAGCCGACCA		3752	Db	2949	LeuAlaGlyGlyValLeuAlaAlaValAlaTyCysLeuAla-----		1680	
Qy	1322	ysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGlnAlaGlu		1342	Qy	1665	CTGGCAGGAGGCCCCCATGGCGACATGGGCTTTTTCACACACTCATGATGTCAACCTCGCA		2835	
Db	3751	CACTTGACGGAAGCCGCGAGATAGTCAGTGAGGTAGCTCCATT-----		3704	Db	2835	-----ThrGlyCysValCysIleIleGlyArgLeu-----		1690	
Qy	1342	hrAlaGlyValArg-----		1355	Qy	1690	GGTTGACCTTGGCAGCTGTGGACAGACACAGGAAACTCTCATAGTCTTCAGACATC		2775	
Db	3703	TGGCAGGGGTTAGGTGAGCACTCGCTCCACCTCGTCCTCAGATGAGTCCCGGTAGCCCT		3644	Db	2775	HisIleAsnGlnArgAlaValAlaProAspLysGluValLeuTyGlnAlaPheAsp		1710	
Qy	1355	roProGlySer-----		1372	Qy	1710	---CTTGATACATCCCTGTTTTCAGGGGTGGTAGCACATGGGGCCCAACACACCTC		2718	
Db	3643	G-CCTGGCAGGCGCAGTAATGTTATCTCCCTCATCCACACA--GAGCCCTCCATGCTGG		3587	Db	2718	GluMetGluCysAlaSerArgAlaAlaLeuIleGluGlyGlnArgIleAlaGlu		1730	
Qy	1372	lyGlnGluGlyGluIleProPheTyArgAlaIleProLeuSerTyIleLysGlyG		1392	Qy	1730	ACAGTGGGTCTCTGTATA-----		2682	

```
QY 1731 MetLeuYsSerTySIIleGInGlyLeuLeuGlnAlaSerTySgInAlaGlnAspIle 1750
    : : : : :
Db 2681 CTTGTATCCAGCGCATC-----ATCAATGCA 2655

QY 1751 GlnProThrValGlnAlaSerTrpProLysValGluGlnPheTrpAlaLysHisMetTrp 1770
    : : : : :
Db 2654 GGTCCCTGGTTGAGGAGG-----2634

QY 1771 AsnPheIleSerGlyIleGlnTyLeuAlaGlyLeuSerThrLeuProGlyAsnProAla 1790
    : : : : :
Db 2633 -----GTTGGAGGCATTGTTGATGTTGTTG-----GCA 2601

QY 1791 ValAlaSerMetMetAlaPheSerAlaAla-----LeuThrSer 1803
    : : : : :
Db 2600 ATTAGGCCCATGAAGCCTTCTCGCAGGTGCATACGCCACTGGTGCATGCTCTTGCA 2541

QY 1804 ProLeuSerThrSerThrThrIleLeuLeuAsnIleLeuGlyGlyTrpLeuAlaSerGln 1823
    : : : : :
Db 2540 GGTGCCACCGTTGACACAGGGTTGGACTCACCCTCGTTGTTGATGTCACAGTTTGT 2481

QY 1824 IleAlaProAlaGlyAlaThrGlyPheValValSerGlyLeuValGlyAlaAlaVal 1843
    : : : : :
Db 2480 TCACCTCCACCGGCGACAGTCACACTTGTACCCATTGAGGCCATCCCGGCAAGCTCC 2421

QY 1844 GlySerIleGlyLeuGlyLysValLeuValAsp-----IleLeuAla 1857
    : : : : :
Db 2420 GTGATGCAGGGGTTACTGTTGCACCTGTTGACTCGGACAGGCACGTGGGGTCATGGTA 2361

QY 1858 GlyTyRAlaGlyIleSerGlyAla-----LeuValAlaPheLysIleMet 1873
    : : : : :
Db 2360 GCCCTCGGGCAGCGGCAAGTGAAGCGCGGATGTCATCTCACAAGTGGCCCGCTTGTG 2301

QY 1874 SerGlyGluLysProSerMetGluAspValValAsnLeuLeuProGlyIleLeuSerPro 1893
    : : : : :
Db 2300 GCAGGG-----GTCGCCGCACATTCGTCATGTTGACGTTACAGTTACACATGCTTCTGT 2250

QY 1894 GlyAlaLeuValValGlyValIleCysAla-AlaIleLeuArgArgHisValGlyProG1 1913
    : : : : :
Db 2249 GTAGCTGTTTCACAGGCACATTCTGTAGCCATCAATCTTCTCCAGACAGGTGCCAGATC 2190

QY 1913 yGluGlyAlaValGlnTrpMetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisVa 1933
    : : : : :
Db 2189 ACAGGGGTTGCTGCGCAGTCATCCAGTTGATC-----TCACAGTTGGGCCCTGT 2139

QY 1933 lAlaProThrHisTyRValThrGluSerAspAlaSerGlnArgValThrGlnLeuLeuG1 1953
    : : : : :
Db 2138 GGTTCCTTGAGGCATAAGCAGAGTAGGAGTTGTCACGG-----2099

QY 1953 ySerLeuThrIleThrSerLeuLeuArgArgLeuHisAsnTrpIleThrGluAspCysPr 1973
    : : : : :
Db 2098 -----TCCTGCAG 2091

QY 1973 oIleProCysGlyGlySer-----TrpLeuArgAspValTr 1985
    : : : : :
Db 2090 GTGCCCCCATGGCGCAGCGTTGGCTGTCGCACTCATTTGATGTTGCTCACAGT-----2036

QY 1985 pAspTrpValCysThrIleLeu---ThrAspPheLysAsnTrpLeuThrSerLysLeuPh 2004
    : : : : :
Db 2035 -GATGGCCTGTAGCCTGGCTGGCCACAGCAGCGTAAAGGTGG-----1994

QY 2004 eProLysMetProGlyLeuProPheValSerCysGlnLysGlyTyRtySgLyValTrpAl 2024
    : : : : :
Db 1993 -----CCACACCATCTTAC-----AGCAACCATAGTGGC-----AG 1962

QY 2024 aGlyThrGlyIleMetThrThrArgCysPro-----CysGlyAlaAs 2038
    : : : : :
Db 1961 GGGTCAGGGT---CACACTCGTCAATGTCCTCACTTCGCGAGTGGCTCCCTGTGTAACTTCT 1905

QY 2038 nIleSerGlyAsnValArgLeuGlySerMetArgIleThrGlyProLysThrCysMetAs 2058
    : : : : :
Db 1904 GTACACACGCGATAGGTGTTGGGCCCATCCAGGCATTCAGCCGTTCTTTCATGTT 1845

QY 2058 nIleTrpGlnGlyThrPheProIle-----2066
    : : : : :
Db 1844 GTGCTGGCAC---ACTCATCCACATCATACTGCGACAGGTGCCGTTGAAGCTTCGGGG 1788

QY 2067 -----AsnCysTyRThrGluGlyGlnCysValProLysProAl 2079
    : : : : :
Db 1787 CACTGACATTGGAACCTCATGATCTTGTCTCATGCGAGTGGCCATTGTGCAGACAGG-----1733

QY 2079 aProAsnPheLysValAlaIleTrpArgValAlaAlaSerGlyTyRAlaGluValThrG1 2099
    : : : : :
Db 1732 -----GGCTGCTGGCGCACTCATCGTGTTCATTTTCACAGTATACACT 1689

QY 2099 nHisGlySerTyRHisTyRileThrGlyLeuThrThrAspAsnLeuLysValProCysG1 2119
    : : : : :
Db 1688 TCATAACCTGGCACACATATGCAATGGA-----1661

QY 2119 nLeuProSerProGluPhePheSerTrpValAspGlyValGlnIleHisArgPheAlaPr 2139
    : : : : :
Db 1661 -----1661

QY 2139 oThrProLysProPhePheArgAspGluValSerPheCysValGlyLeuAsnSerPheVa 2159
    : : : : :
Db 1660 -ACTCCCCAATC-TGGTCCAGCAAGTGGCATCATTTGCACATGGTGTGAGATGCACCTC 1603

QY 2159 lVal-----GlySerGlnLeuProCys-----AspProG1 2169
    : : : : :
Db 1602 ATTAACATCAATCTCACAGCGGGTCCCGTGTAGCCCTGTAGACACTGGCACCTCAAAAGA 1543

QY 2169 uProAspThrAspValLeuMetSerMetLeuThrAspProSerHisIleThrAlaGluTh 2189
    : : : : :
Db 1542 ACCC-----AGTGTGTTG-----AGGCATTTTCCCTGCG-----1515

QY 2189 rAlaAlaArgArgLeuAlaArgLysSerProProSerGluAlaSerSerAlaSerG1 2209
    : : : : :
Db 1514 -----TGCTCAAGGGTGGCACCC---AGAGCACACTCATCCAGTCTCTG 1471

QY 2209 nLeuSerAlaProSerLeuArgAlaThrCysThrThrHisGlyLysAlaTyRAspValAs 2229
    : : : : :
Db 1470 GCTGCAGGCT-----GGCCCTGTGTACCCCGAGGG 1441

QY 2229 pMetValAspAlaAsnLeuPheMetGlyGlyAspValThrArgIleGluSerGlySerly 2249
    : : : : :
Db 1440 GCAGGTGCAGATGGCTTTTCCCGTTGCAGAGGTGTTG-----1404

QY 2249 sValValValLeuAspSerLeuAspProMetValGluGluArgSerAspLeuGluProSe 2269
    : : : : :
Db 1403 -----TCACGTTGGAGCCCTC 1387

QY 2269 rIleProSerGluTyRMetLeuProLysLysArgPheProProAlaLeuProAlaTrpAl 2289
    : : : : :
Db 1386 GTTGACAGGGTGTCTGATGCACGCATCGTTGAG-----TGGCA 1348

QY 2289 aArgProAspTyRAsnPro-----ProLeuValGluSerTrpLy 2302
    : : : : :
Db 1347 CACGACAGCCTGTGCGCCCATTCGCGCATTCGCGAGTAGAAGGAAGGCACACGGTCTGGCA 1288

QY 2302 sArgProAspTyRTrpProAlaThrValAlaGlyCysAlaLeuProProProArgLysTh 2322
    : : : : :
Db 1287 AGTGACCCCTGGAAACAGCGCGCACTGGCACAGTCATCATGTTCTCTACTGAGCTCTC 1228

QY 2322 rProThrProPro-----Pr 2327
    : : : : :
Db 1227 GCCAGTCCACCCATTGACACACACACACACTGTTGAGCGCGGTGTGTGTCGAGGTTCC 1168

QY 2327 oArgArgArgArgThrValGlyLeuSerGluAspSerIleGlyAspAlaLeuGlnGlnLe 2347
    : : : : :
Db 1167 GCCATTCTGGCAGGCAATGGGCATGAGTGCATTCGTCACATCTCTCTGTACAGTACTG 1108

QY 2347 uAlaIleLysSerPheGlyGln-----Pr 2355
    : : : : :
Db 1107 ACCCGTTCACCTCCGTGGCGCAGCGCAATTTAGTAGTTATTCAGCGGTCCACACAGCAC 1048

QY 2355 oPro-----ProSerGlyAspSerGlyLeuSerThrGlyAlaGlyAlaAl 2370
```

Db 1047 CCATTCTCCAGTGTTCCTCGACGATCAATCCATTTCTTACAGCTTGTCTCCAC 988
 Qy 2370 aaspSerGlySerGlnThrProProaspGluLeuAlaLeuSerGluThrGlySerIleSe 2390
 Db 987 AAAACCTGGCAAGCAGCA-----CACTCGTGGTGTCTCCCTGTAGGA-----CG 940
 Qy 2390 rSerMetProPro-----LeuGluGlyGluLeuGluGlyAspProaspIle 2404
 Db 939 GCAGGTGCTCCATTTCTGCGAGGTGAGGGCTGCAGGGCAGATAGGGCAGTTCACAGTG 880
 Qy 2404 uGluProGluGlnVal-----GluPro-----GlnPr 2413
 Db 879 GGCACCATGATGGTGGCCAGCGGAGGACAGCAGCATAGAGCCGATCTCATTTGGGAGGT 820
 Qy 2413 oProPro 2415
 Db 819 GCCTCCA 813

RESULT 5
 CNSLTI164
 LOCUS
 DEFINITION human full-length cDNA clone CSODM003YJ21 of Fetal liver of Homo sapiens (human).
 ACCESSION BX248258
 VERSION BX248258.1 GI:28207864
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 2891)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished
 Contact: Feng Liang Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue
 2 (bases 1 to 2891)
 Genoscope.
 Direct Submission
 Submitted (04-FEB-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.
 Location/Qualifiers
 1..2891
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSODM003YJ21"
 /tissue_type="Fetal liver"
 /plasmid="pCMVSPORT-6"
 <1..>2891
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD62586.1"
 /db_xref="GI:28207865"

FEATURES
 source
 1..2891
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSODM003YJ21"
 /tissue_type="Fetal liver"
 /plasmid="pCMVSPORT-6"
 <1..>2891
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD62586.1"
 /db_xref="GI:28207865"

CDS
 1047 CCATTCTCCAGTGTTCCTCGACGATCAATCCATTTCTTACAGCTTGTCTCCAC 988
 2370 aaspSerGlySerGlnThrProProaspGluLeuAlaLeuSerGluThrGlySerIleSe 2390
 987 AAAACCTGGCAAGCAGCA-----CACTCGTGGTGTCTCCCTGTAGGA-----CG 940
 2390 rSerMetProPro-----LeuGluGlyGluLeuGluGlyAspProaspIle 2404
 939 GCAGGTGCTCCATTTCTGCGAGGTGAGGGCTGCAGGGCAGATAGGGCAGTTCACAGTG 880
 2404 uGluProGluGlnVal-----GluPro-----GlnPr 2413
 879 GGCACCATGATGGTGGCCAGCGGAGGACAGCAGCATAGAGCCGATCTCATTTGGGAGGT 820
 2413 oProPro 2415
 819 GCCTCCA 813

BASE COUNT 590 a 1003 c 822 g 476 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0.009 Length: 2891
 Score: 185.00 Matches: 209
 Percent Similarity: 31.67% Conservative: 120
 Best Local Similarity: 20.12% Mismatches: 360
 Query Match: 1.15% Indels: 353
 DB: 11 Gaps: 44
 US-09-980-559-2 (1-3033) x CNSLTI164 (1-2891)
 Qy 1912 ProGlyGluGlyAlaValGlnTrpMetAsnArgLeuIleAlaPheAlaSerArgGlyAsn 1931
 Db 115 CCCCGTCCGAGCTGCCGCGGCATGATCCGACACGCGCGGGCGCCGCGCGGGAC 174
 Qy 1932 HisValAlaProThrHisTyrValThrGluSerAspAlaSerGlnArgValThrGlnLeu 1951
 Db 175 CCCACGGGTCCGTTCCAGTTCTGCGCAAGAGAGAGAGAGAGAGAGATGGCATG 234
 Qy 1952 LeuGlySerLeuThrIleThr-----SerLeu 1960
 Db 235 CGGCTTTGTCTCCACCAACCCGAAACTGCTCTCTCCCGCGGGGCATCAGCATC 294
 Qy 1961 LeuArgGluLeuHisAsnTrpIleThrGluAspCysProIle-ProCysGlyGlySerTr 1980
 Db 295 CTGGAGAGCTC-----ATCAAAACATCCCGGTGTGCTGCGAGTGTGCTG 342
 Qy 1980 pLeuArgAspValTrpAspTrpValCysThrIleLeuThrAspPheLysAsnTrpLeuTh 2000
 Db 343 GCCAGGAGAGGTGCCAGGATCTGCACC-----GGGTGG----- 379
 Qy 2000 rSerLysLeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTyrL 2020
 Db 380 -----GGTGGGATGTTC----- 394
 Qy 2020 sGlyValTrp-----AlaGlyThrGlyIleMetThrArgCysProCysGlyAlaAsnI 2039
 Db 395 -----TGGTTCGCGGAGCAGCTCGAAGCAGCTGCTGCTGTCTGCTCCACTTCT 447
 Qy 2039 e-----SerGlyAsnValArgLeu----- 2045
 Db 448 TCTCTGAACGAAAGCTCGCGCGAGGTGCTCGAATACACCATTAAGGAAGAAAGTCGATA 507
 Qy 2046 -----GlySerMetArgIleThrGlyProLysThrCysMetAsnIleTrpGlnGlyTh 2063
 Db 508 TTGTACCTCGAAGGCTCGGCTCTTGTGTTTGGAGCATCTTCAGATGATGTTGTTCTAC 567
 Qy 2063 rPheProIleAsnCysTyrThrGluGlyGlnCys-----ValProLysProAlaProAsnPh 2082
 Db 568 TGTGTGATGAGACACTTACTGCGCTTCACACTGGGCTACCCAGGCGCATCC----- 619
 Qy 2082 eLysValAlaIleTrpArgValAlaAlaSerGluTyrAlaGluValThrGlnHisGlySe 2102
 Db 620 -----TTGAGGCGCAGCAGCTTCACGACCTTGAGA----- 649
 Qy 2102 rTyrHisTyrIleThrGlyLeuThrThrAspAsnLeuLysValProCysGlnLeuProSe 2122
 Db 650 -----CCATC----- 654
 Qy 2122 rProGluPhePheSerTrpValAspGlyValGln-----IleHisArg----- 2136
 Db 655 GCCAACCTGGGTC---TGGGTTCTGGGACTCTCTCGTCTGATCTCCACAGAAAGAGGG 711
 Qy 2137 -----PheAl 2138

```
Db 712 AAGCAGCAGAGCCCAAGAGAGCCGGGCCCCCGGATTCCCTAGTCTCCAGCCTCAGG 771
Qy 2138 aProThrPro-----LysProPheArgAspGluValSerPh 2151
Db 772 CCCACAGCCCATGACGCAAACTGCTGTGAATCGAGCTGTCGGTAGAATGACCGC 831
Qy 2151 eCysValGlyLeuAsnSerPheValVal-----GlySerGlnLeuProCysAspPr 2168
Db 832 CTGTGG-TTGTGAATCTATTTCATCGAGGACTGCAGCAGCGCCCTGCCACCGACCA 890
Qy 2168 oGluProAspThrAspValLeuMetSerMetLeuThrAspProSerHisIleThrAlaGl 2188
Db 891 GCCACCTCTTTGGAAATGGCCCTGACGCCCTTTGGCCGCCACCTCTGATGCCACCTCACC 950
Qy 2188 uThrAlaAlaArgArgLeuAlaArgGlySerProProSerGluAlaSerSerAlaSe 2208
Db 951 CACCTCC-----AGGTGGGCCCCCAGCCGCCCCACACACCCCTCCCCAGTGTGCCCT 1004
Qy 2208 rGlnLeuSerAlaProSerLeuArgAlaThrCysThrThrHisGlyLysAlaTyrAspVa 2228
Db 1005 GCAGCCCTGCAGCCCA----- 1020
Qy 2228 lAspMetValAspAlaAsnLeuPheMetGlyGlyAspValThrArgIleGluSerGlySe 2248
Db 1021 -----GC 1022
Qy 2248 rLysValValValLeuAspSerLeuAspProMetValGluGluArgSerAspLeuGluPr 2268
Db 1023 CCAGCCCCCTGCTCTCTCTGCCCCCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1082
Qy 2268 oSerIleProSerGluTyrMetLeuProTyrLysArgPheProProAlaLeuProAlaTr 2288
Db 1083 CCCAGTGCCTGCCCCACGCTCACACCCATGCCCGCCAGGTCCCCAGACCATCG----- 1137
Qy 2288 pAlaArgProAspTyrAsnProProLeuValGluSerTrpLysArg----- 2303
Db 1138 -----AACCAGCGCCCATGATGACCTGCGAGAGACTCCCATGCCCCAC 1181
Qy 2304 -----ProAspTyrGlnProAlaThrValAlaGlyCysAla----- 2315
Db 1182 TGCAGGCTGGGCCCCCTCAGGAGGAGGAGCGATGAGCGAGGCGAGCTTCCAGTCCCTT 1241
Qy 2316 -----LeuProProProArgGlyThrProThrProProPro 2328
Db 1242 GCAGCAGTCCCGCCCGCCACTGCTGCGAAGAGAACCTTCCCACTGCCCTTCCAG 1301
Qy 2328 gArgArg-----ArgThrValGlyLeuSerGluAspSerIleGlyAspAlaLeuGln-- 2345
Db 1302 ACGCCGCTTTCGAGAGGGTGTCTTAGAAGACCAAAAGTCCGGGGATGGCGCAGAGGG 1361
Qy 2346 -----GlnLeuAlaIleLysSerPheGly-----GlnProProPr 2357
Db 1362 GGACAGCTCAGCTGCTCCCAAGGAGACTCAGAGCGGCCCTGAGACACGCCCCGGGA 1421
Qy 2357 oSerGlyAspSerGlyLeuSerThrGlyAlaGlyAlaAlaAspSerGlySerGln----- 2375
Db 1422 GAGCAGGAGCAAGCCAGGACACAGAGGTGAAGCCAGCGATCCTCAGCAGCATGCCAGA 1481
Qy 2376 -----ThrProProAspGluLeuAlaLeuSerGl 2385
Db 1482 GCTGCCCAGGACAGCCAAACACACCCCGCAGTCCCGCCGCCCGCCAGGAAACCGATCTCTCG 1541
Qy 2385 uThrGlySerIleSerSerMetProProLeuGlu-----GlyGluLeuGlyAspProAspLe 2404
Db 1542 ACAACTGGCTCGACCCCTCCAGCTCCTTAGAGAACGCTGAGCTGTGCACACAGCGCAT 1601
Qy 2404 uGluProGluGlnValGluProGlnProProGlnGly----- 2417
Db 1602 GGCTTTGGAGACACCCACCGCGGGTCCACCCAGAGAGGGCCAAAGCCCTGCTTCTCAGGC 1661
Qy 2417 ----- 2417
Db 1662 TGGGACTCAGCACCCCTCTGCCCAGGCCACTGCCCATTTCCAGAGCTCTCCAGAGTTCAA 1721
```

QY	2718	-----TyArgArgCysArg-----	-----AlaserGlyValLeuHth 2728
	2677	CGCAGCTGAGTGGAGTGGAGGACTCATCCAGCGTGGAGCGCGGCTACTCTC 2736	
QY	2728	rThrSerMetGlyAsnThrIleThrCysTyrVallyAlaLeuAlaCysLys 2746	
	2737	AACAAGCGCGGCGCTCCCGCTCCTCCGTAC---AGGACTTCATCTGCGTGTGT 2788	
RESULT 6			
LOCUS	AK032205		
DEFINITION	Mus musculus adult male olfactory brain cDNA, RIKEN full-length enriched library, clone:6430500K07 product:similar to RIN3 (FRAGMENT) (Homo sapiens), full insert sequence.		
ACCESSION	AK032205	3865 bp	HTC 05-DEC-2002
VERSION	AK032205.1	GI:26328032	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Carninci, P. and Hayashizaki, Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20499374		
PUBMED	11042159		
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishimura, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakauchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		
MEDLINE	20530913		
PUBMED	11076861		
AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Giesi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombarts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohse, S. and Hayashizaki, Y.		
TITLE	Functional annotation of a full-length mouse cDNA collection		
JOURNAL	Nature 409 (6821), 685-690 (2001)		
MEDLINE	21085660		
PUBMED	11217851		
REFERENCE	5		
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.		
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs		
JOURNAL	Nature 420, 563-573 (2002)		
REFERENCE	6 (bases 1 to 3865)		
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, M., Hiramoto, K., Hiraoka, T., Hirozane, T., Kori, F., Motani, K., Iehii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koyama, S., Kurthara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)		
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.		
FEATURES	Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/ Location/Qualifiers 1..3865 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM DB:6430500K07" /db_xref="taxon:10090" /clone="6430500K07" /sex="male" /tissue type="olfactory brain" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="adult" 214..3156 /note="unnamed protein product; putative similar to RIN3 (FRAGMENT) (Homo sapiens) (SPTP BAB84316, evidence: PASTY, 84.8%ID, 100%length, match=1563)" /codon_start=1 /protein_id="BAC27757.1" /db_xref="GI:26328033" /translation="MRAEAPSSAPHPGIPDAGKEGEDEEEDKGTGLSTTPRNC IPRGGISVLEKLVKPKESKAILYLGSLVRFEDILRIAFYCVSRDLPTLRLOAI PSLKGSSAEVLVPTKEKAILYLGSLVRFEDILRIAFYCVSRDLPTLRLOAI LEASFELEETISNLGLFWDSLSNRGSAEPLSPACPTASSSLRPPTHVANCSC IELSVGNRLWFNFIFEDICLPADPPPLPGTSPPTPTPATPATSTSTSGSPRP PPPLPTVPTGPARLAPVPVPPHQAQPPPLPKKALPAAPRRISERVSLES ESLPFAVGLGPGGEEKMPTNPLHQAQPPPLPKKALPAAPRRISERVSLES QNVGTSRDRHSGISRTASLNLPQSTVSSLDGPRPTTEQSDTEAKASHADSI PVP PKAKQPVPPPKKRVSRQLASTLLSPLESPIQASSEKQATGASWELGSPVQAGM QHLQVSSSCQSSPEFGSOASLSGLGVPASADODSYSTSSAEFELEFTSPDVEK KESMLDDKARHLRSFVSFAVFAFLSDRKLKVKVELAQDKSSYFGSLVDQYKYS LEMAQRTSSLEIHTSRDGLQKSLQSTELKALVEPTLHSEELSAI VESALYK CVLXPKMGKAYSPEKKISILLTKCLYDSMALGNPKGPKGADDFLPLMYLVARSLNLT OKLTMHKAYSPEKKISILLTKCLYDSMALGNPKGPKGADDFLPLMYLVARSLNLT EMLNVEYMWELDPALOGESYLLTYGALEHIKNYDKITVTRQLSVQVDSIHR WERRTLNKARASRSVODFCVSVLKEPQSQSTLASRADTAQAALCAACAQFEVSO POYRLFLVLDGRCFQLADEALPHRIKNGYLLRSEPKRDFHVFYRQDSKQDASSQFCI VVREPNFL" 3842..3847 /note="putative"		
polyA_signal	3842..3847		

Copied from 09540843 on 05/19/2004

2646	LeuLeuLysAlaTrpAlaGluLysLysAspProMet-----GlyPheSer	2660
2488	TTGTGAAGACCTGCAAGCTCATCTATGACTCCATGCTTGGCAACCAGGGAAGCCC	2547
2661	TyrAspThrArgCysPhe-----AspSerThrValThr	2671
2548	TACGGGGCGGACGACTTCCTGCCTGTGCTCATGTACGTGTGGCCGCCGAGCAACCTCACT	2607
2672	GluArgAspIleArgThrGlu-----GluSerIleTyrArgAlaCysSerLeuPro	2688
2608	GAGATGCTCTCCAACGCTGAGTACATGATGGAGCTCATGGACCCCGCCTGCAGTTAGA	2667
2689	GluGluAlaHis-----ThrAlaIleHisSerLeuThrGluArgLeuTyrValGlyGly	2706
2688	GAGGGTCTCTACTATCTGACCACCACTCACGGAGCCCTGGAGCACAATTAAA-----	2718
2707	ProMetPheAenSerLysGlyGlnThrCysGlyTyrArgGlyCysArgAlaSerGlyVal	2726
2719	-----AACTATGACAAGATCACAGTCACAGTGACCCGACAG	2748
2727	LeuThrThrSerMetGlyAenThrIleThrCystYrValLysAlaLeuAlaCysLys	2746
2749	CTGACGGTGGAGGTGACGAGCTCCATCTGCTGGAAACGCAG-CGCGACGCTCAAACA	2807
2747	AlaAlaGlyIleIleAlaProThrMetIleValCysGlyAspAspLeuValIleSer	2766
2808	AGCGGGGGCCTC-----CCGTCCTCTGTGCGAGGACTTCATCTGTGTCTCTACCT	2858
2767	GluSerGlnGlyThrGluGluAspGluArgAenLeuArgAlaPheThrGluAlaMetThr	2786
2859	GAAGCCCGACGACGATCAGGAC-----ACTGGCGTCACGGGC	2897
2787	ArgTyrSerAlaProProGlyAspProProArgProGluTyrAspLeuGluLeuThr	2806
2898	AGACACAGCAGCCCCAGGC-----ACT	2918
2807	SerCysSerSerAsnValSerVal-----	2814
2919	GTGTGCACAGTCGCGTGAAGACTTTGAGGTGTCAAGCCCCAAGACTACCGGCTCTTCGT	2978
2815	-----AlaLeuGlyProGlnGlyArgArg	2822
2979	GCTGTGGACGGCGCTGTTCCAGCTGGCCGACGA	3014
<hr/>		
RESULT 7		
AK004685	2861 bp mRNA linear HTC 05-DEC-2002	
LOCUS	Mus musculus adult male lung cDNA, RIKEN full-length enriched	
DEFINITION	library, clone:120010B10 product:CAPICUA PROTEIN homolog.[Mus	
	musculus], full insert sequence.	
ACCESSION	AK004685	
VERSION	AK004685.1 GI:12836036	
KEYWORDS	HTC; CAP trapper.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;	
REFERENCE	1	
AUTHORS	Carninci, P., Shibata, Y., Hayateu, N., Sugahara, Y., Shibata, K.,	
TITLE	Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	
JOURNAL	Normalizing and subtraction of cap-trapper-selected cDNAs to	
MEDLINE	prepare full-length cDNA libraries for rapid discovery of new genes	
PUBMED	Genome Res. 10 (10), 1617-1630 (2000)	
REFERENCE	2	
AUTHORS	10349636	
TITLE		
JOURNAL		
MEDLINE		
PUBMED	20499374	
REFERENCE	3	
AUTHORS	11042159	

FEATURES	Location/Qualifiers
source	1. .2861

CDS

```
/codon_start=2
/protein_id="BAB23472.1"
```

polyA signal

polya_signal	2861
polya_site	2861
/note="putative"	2861

	COUNT	576 a	987 c	729 g	569 t
/note="putative"					

100

ment Scores:

No.:	0.0231	Length:	2861
:	179.50	Matches:	202
nt Similarity:	31.89%	Conservative:	107
Local Similarity:	20.95%	Mismatches:	375
Match:	1.11%	Indels:	285
	11	Gaps:	48

US-09-980-559-2 (1-3033) x AK004685 (1-2861)

Qy 2008 ProGlyLeuProPheValSerCysGlnLysGlyTyrLysGlyValTrpAlaGlyThrGly 2027

207 CCCTGGTGCCTCTAGCATGTCAGTTCGAGGTGAGGGG----- 245

2028 I l e m e t h r T h r A r g C y s P r o C y s G l y A l a A s n I l e S e r - G l y A s n V a l A r g L e u G l y S e 2047

246 ---CTGTCAGCCACTGCCCTCTGGTTAGCTCGCCCTTCTCAGT--- 285

2047 rMetArgIleThrGlyProIysThrCysMetAsnIle----- 2059

286 -----ACCTGTCCAAAATGTTGCCCAACAACCTAGCAAGATTATCCAGTGA CTC 335

QY 2060 -----TrpClnGlyThrPheProIleAsnCysTyrThrGluG1 2072

336 CTGTGCTGTGAGCACACCTAGTGGCTGGTACCAACCTGAGCCCAAGTACAAATGCCGG 395

2072	Y	Gln	Cys	Val	Pro	lys	-----	Pro	Ala	Pro	Asn	Phe	-----	-----	Ly	2083			
	---	---	---	---	---	---	---	---	---	---	---	---	---						
396	G	A	C	C	A	C	A	G	C	C	T	C	A	G	A	455			
	---	---	---	---	---	---	---	---	---	---	---	---	---						
2083	s	Val	Ala	ile	Trp	Arg	Val	Ala	A	As	e	G	lu	Tyr	Ala	2100			
	---	---	---	---	---	---	---	---	---	---	---	---	---						
456	A	G	T	C	A	G	A	G	T	G	G	C	A	C	A	515			
	---	---	---	---	---	---	---	---	---	---	---	---	---						
2101	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----						
	---	---	---	---	---	---	---	---	---	---	---	---	---						
516	C	A	G	T	G	A	C	C	T	C	T	A	T	G	G	575			
	---	---	---	---	---	---	---	---	---	---	---	---	---						
2110	r	Th	r	As	p	As	n	L	e	u	lys	Val	Pro	Cys	Gln	2129			
	---	---	---	---	---	---	---	---	---	---	---	---	---						
576	G	T	G	E	T	C	T	C	T	T	T	A	C	A	G	635			
	---	---	---	---	---	---	---	---	---	---	---	---	---						
2129	l	As	p	-	G	ly	V	al	G	ln	ile	H	i	s	Arg	2149			
	---	---	---	---	---	---	---	---	---	---	---	---	---						
636	A	G	T	G	G	G	-	G	T	G	-----	-----	-----	-----					
	---	---	---	---	---	---	---	---	---	---	---	---	---						
2149	a	l	S	e	r	P	h	e	C	ys	Val	G	ly	-----	-----				
	---	---	---	---	---	---	---	---	---	---	---	---	---						
681	G	T	A	C	A	G	C	C	T	C	T	G	T	C	A	740			
	---	---	---	---	---	---	---	---	---	---	---	---	---						
2167	s	P	ro	G	lu	-----	-----	-----	-----	-----	-----	-----	-----						
	---	---	---	---	---	---	---	---	---	---	---	---	---						
741	C	A	G	A	C	C	C	T	G	G	G	C	C	A	C	800			
	---	---	---	---	---	---	---	---	---	---	---	---	---						
2185	l	e	Th	r	Ala	G	lu	T	h	r	Ala	a	l	a	Arg	2205			
	---	---	---	---	---	---	---	---	---	---	---	---	---						
801	G	T	A	C	A	C	T	C	C	A	G	T	C	T	C	860			
	---	---	---	---	---	---	---	---	---	---	---	---	---						
2205	e	r	S	e	r	Ala	S	e	r	G	ln	L	e	u	S	e	r	Ala	2225
	---	---	---	---	---	---	---	---	---	---	---	---	---						
861	C	T	C	A	G	C	C	C	T	A	G	T	C	T	T	C	C	907	
	---	---	---	---	---	---	---	---	---	---	---	---	---						
2225	l	a	T	r	As	p	Val	As	p	Met	Val	As	p	Ala	Asn	L	e	u	2245
	---	---	---	---	---	---	---	---	---	---	---	---	---						
908	-	-	T	A	C	A	G	T	T	A	G	T	G	G	T	C	C	965	
	---	---	---	---	---	---	---	---	---	---	---	---	---						
2245	l	u	S	e	r	G	ly	S	e	r	Val	Val	L	e	u	As	p	2265	
	---	---	---	---	---	---	---	---	---	---	---	---	---						
966	A	G	G	C	A	G	C	A	T	T	C	T	G	T	G	G	G	1025</	

Copied from 09540843 on 05/19/2004

2193 Db -----TGTTCCCGAGGCCCTCGCCTTGTGCTATGG 2224
 1605 QY ---ProSerThrAspValMetTrpLysCysLeu-----ThrArgLeuLys-ProT 1621
 2225 Db CGCTGTGAGAACACGAGGAGCTTCAAGTGTCTGCTGCCCAACAGGCTTCCACCCAA 2284
 1621 QY hrLeuValGlyProThrProLeuLeuTyArgLeuGlySerValThrAsnGluValThrL 1641
 2285 Db CGCTGGCGGCTCCGAGTGCAGAGTGG----- 2313
 1641 QY euThrHisProValThrLysTyIleAlaThrCysMetGlnAlaAspLeuGluValMetT 1661
 2314 Db -----ATGAGTGTGAGAA 2326
 1661 QY hrSerThrTrpValLeuAlaGlyGlyValLeu-----AlaAlaValAlaAlaT 1677
 2327 Db ---TGCGTGTCCAGGCGAGGAGTGTGTGAACCTACCGGGCTCTTCCAGTGGCG 2383
 1677 QY yrCysLeuAlaThrGlyCysValCysIleleGlyArgLeuHisIleAsnGlnArgAlav 1697
 2384 Db GGCCTGCCCTGTGGGACCACTGACCGCTGGCAGATGTACTGTGGATGAATGCCAG 2443
 1697 QY alValAlaProAspLysGluValLeuTyArgLeuAlaPheAspGluMetGluGluCysAlas 1717
 2444 Db TTGCGGACCC-----CTTGTGTCTCCATGGCCAGTGCACAAATACTAAAGGCTCTT 2497
 1717 QY erArgAlaAlaLeuIleGlu---GluGlyGlnArgIleAlaGluMetLeuLysSerLysI 1736
 2498 Db CCACCTGCAGCTGCTCGACAGATTACCGGGCGCATCTGTGTCAACCTG----- 2544
 1736 QY leGlnGlyLeuLeuGln-----GlnAlaSerLysGlnAlaGlnAspIleGlnProThrV 1754
 2545 Db ---GGCCTGTGTGCACACATAAATGAGTGTCTGGAAGGCGACTTCTGCTTCCCGCCACGG 2599
 1754 QY alGlnAlaSer-----Trp 1759
 2600 Db AGAGTGCTTACACTGATGGCTCTTACTGTACCTGTGCTGCTGCTGCTGCTGCTGCTG 2659
 1759 QY roLysValGluGlnPheTrpAlaLysHisMetTrpAsnPhelSerGlyIleGlnTyL 1779
 2660 Db ACCTCGTGGAGCATCTT-----GTCGTGAGCTGGAGCTGAGTGCAGCGAAGAGGA 2707
 1779 QY eu---AlaGlyLeuSerThrLeuProGlyAsnProAlaValAlaSerMetMetAlaPheS 1798
 2708 Db TCCTTGGCAGAGCGGCATCTGTACCA----- 2733
 1798 QY erAlaAlaLeuThrSerProLeuSerThrSerThrThrLeuLeuAsnIleLeuGlyG 1818
 2734 Db ---ACACTGACGGCTCTTCGAGTGTGATCTGTC-----CTCCTGGACA 2773
 1818 QY lyTrpLeuAlaSerGlnIleAlaProProAlaGlyAlaThrGlyPheValValSerGlyL 1838
 2774 Db CGCGCTGGCCCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2833
 1838 QY euValGlyAlaAlaValGlySerIleGlyLeuGlyLysValLeuValAspIleLeuAlaG 1858
 2834 Db TGCCCTGTGCGGGTCTGACGCTGTG-----AAATCTCTGCTGCTTCTTACCCTG 2884
 1858 QY lyTyGlyAlaGlyIleSerGlyAlaLeuValAlaPheLys-----IleM 1873
 2885 Db TGTGGGAGCTGCGATCCAGGTTACCCCGGCTGAGGCGCTGCTGATGATGATGA 2944
 1873 QY etSerGlyGluLysProSerMetGluAspValValAlaAsnLeuProGlyIleLeuSerP 1893
 2945 Db TGAGTGCCGAGAAATACGCTCTGCGATTTGTGGTG-----CCACGCGCTGTGAGAA 2995
 1893 QY roGlyAlaLeuValValGlyValIleCysAlaAlaIleLeuArgHisValGlyProG 1913
 2996 Db CACCCCGGCTCTTACCCTGACACACGCTGCGACCTGCTGCTATCATCACCACAGG 3055
 1913 QY lyGluGlyAlaValGlnTrpMetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisV 1933
 3056 Db GGGCGGGTGCCAGATGTGGATGAAT-----GCCGGAACCGGCTCTTCTG 3100

1933 QY aLalaproThrHisTyValThrGluSerAspAlaSerGlnArgValThrGlnLeuLeuG 1953
 3101 Db CGGCGCCCATG-----CTATGTGCCAAGATCTGCTGGTGGTCTT 3139
 1953 QY lySerLeuThrIleThrSerLeuLeuArg-----ArgLeuHisAsnTrpIleT 1969
 3140 Db CCAGTGGCTGTGTGACCAAGGCTACGAAGGGGACGGGACGACGCTCACTGGGTGGATGT 3199
 1969 QY hrGluAspProIleProCysGlyGlySerTrpLeuArgAspVal----- 1984
 3200 Db G-AATGAGTGTGAGAC-ACTACAGGGCGTGTGTGCTCCGCTGTGTGAAAATGTGAA 3257
 1985 QY -----TrpAspTrp---V 1988
 3258 Db GGTCTCTTCTGTGTGTGCTGCCCCAATAGCCCTGAGGAATTTGATCCCATGACTGGACGC 3317
 1988 QY alCysThrIleLeuThrAspPheLysAsn-----TrpLeu----- 1999
 3318 Db TGTGTCTCCCGACGGCTCTCGAGGCAGCTTCCCTGGCTCCAGCCCAAGCACCTGCC 3377
 2000 QY -----ThrSerLysLeuPheProLysM 2007
 3378 Db AGCCCGAGTCTTCCAGCCAGGCCACCGACCCCTCCCGACCTCGAAGACCCAGCCCTCC- 3436
 2007 QY etProGlyLeuProPheValSerCysGlnLysGlyTyLysGlyValTrpAlaGlyThrG 2027
 3437 Db -----CAGCAGGGTCCCGTGAGCAGTGGAGCTGCCGAG 3470
 2027 QY lyIleMet-ThrThrArgCysPro-----CysGlyAlaAsnIleSerGlyAsnValArg 2044
 3471 Db TGCTACTTTGATACAGCCCTCCAGATGCTGTGACAATATCTTGGCCCGAAACGTG--- 3527
 2045 QY LeuGlySerMetArgIleThrGlyProLysThrCysMetAsnIleTrpGlnGlyThrPhe 2064
 3528 Db ---ACGTGGCAGGAGTGTCTGTACTGTGGGAGAGGC----- 3563
 2065 QY ProIleAsnCysTyThrGluGlyGlnCysValProLysProAlaProAsnPheLysVal 2084
 3564 Db -----TGGGGAGCGGC---TGGCGCATCCACAGTCCCGGCGCCAGCAGACA 3608
 2085 QY AlaIleTrpArgValAlaAlaSerGluTyAlaGluValThrGlnHisGlySerTyHis 2104
 3609 Db GCT-----GAATACCACTCATTTGTCTCTCACGCCCGGGGTAC 3647
 2105 QY TyIleThrGly-----LeuThrThrAspAsnLeuLysValProCysGlnLeu----- 2120
 3648 Db CTGCTGCCAGTGGAGACCTGAGTGGCCGAGAGATGTGGACGAATGCCCACTTCTCCAG 3707
 2121 QY -----ProSerProGluPhePheSerTrpVal 2129
 3708 Db GACAGGTGTGAAGAGCGGAGTGTGTGTAACACGCGCCCGAGGCTACTCTTGTATTGC 3767
 2130 QY AspGly-----ValGlnIleHisArgPheAlaProThrProLysProPhePheArg 2146
 3768 Db AGCAACGGTCTTACTATCATGCACACAGCTGGAGTGGCTTGTATAATACGAATGTGCG 3827
 2147 QY AspGluValSerPheCysValGly-----LeuAsnSerPheValValGlySer--- 2162
 3828 Db GACGAGGAGCTGTCTTGAAGGTGGCGCTGCTGCTCAACACG-----GTGGGGCTCTTAC 3881
 2163 QY GlnLeuProCysAspProGluProAspThrAspValLeuMetSerMetLeuThrAspPro 2182
 3882 Db CACTGTACCTGCCAGGCC-----CCA 3902
 2183 QY SerHisIleThrAlaGluThrAlaAlaArgLeuAlaArgGlySerProProSerGlu 2202
 3903 Db CTGGTGTGGAGCGCTCCCG-GCGACGCTGCTCTCCAAAGCAGAGCAGGCTGGATGA 3961
 2203 QY AlaserSerAla-----SerGlnLeu-SerAlaProSerLe 2215
 3962 Db CAATCTAGGAGTGTGCTGGCAGGAAGTGGGACCTGATCTCTGTCTGCAGTCCGCCACGACT 4021

2215 u-----ArgAlaThr-----CysThrThrHisGlyLysAlaTyrAspValAs 2229
 Db AGACCGGAGCCACCTACACCGAGTCTCTCTATGGTGAAGCCTGGGGTATGGA 4081
 2229 pMetValAspAlaasnLeuPheMetGlyGlyAspValThrArgIleGluSerGlySerLy 2249
 Db CTGC----- 4085
 2249 sValValValLeuAspSerLeuAspProMetValGluGluArgSerAspLeuGlu----- 2267
 Db -----GCTCTCTCTCT-----GCCAGGACTCAGATGACTTTGAGGCCCT 4126
 2268 -----ProSerIleProSerGluTyrMetLeuPr 2277
 Db TTGCAATGTCTGCTGCTCACCTGCTATGCGCGCGCGCTGCTGGTCTTGGAAATTC 4186
 2277 oLysLysArgPheProProAlaLeuProAlaTyrAlaArgProAspTyrAsnProProLe 2297
 Db CTATGAATATGGCCAGATATAGCCACCTTATCAGAGTCTCCCTATGGCCAGATCT 4246
 2297 uValGluSerTrpLysArgProAspTyrGlnProAlaThrValAlaGlyCysAlaLeuPr 2317
 Db -----TATCACCAGCAGTGTACCTACGCCCTACCC 4282
 2317 oProPro----- 2319
 Db ACCCCACCTGGACCTTTGCTCGCGGAGGCGCTTATGGCGCCCGCTTCGCAT 4342
 2320 -----ArgLysThrProThrProProPr 2327
 Db GCCGGAATTTGAGGATGATGGTGGCCCTATGGCGAGTCTGAGACCTCTGACCCACTAG 4402
 2327 oArgArgArgThrValGlyLeuSerGluAspSerIleGlyAspAlaLeuGlnLe 2347
 Db CCAGGCACGTGTGGCTTATCGTCCAGACACCCGTGCG----- 4445
 2347 uAlaIleLysSerPheGlyGlnProProProSerGlyAspSerGlyLeuSerThrGlyAl 2367
 Db TCCTTCCAGAGCCTGAGGAATCTCTGAGCGTGAAGCTATACAGGA-- 4493
 2367 aGlyAlaAlaAspSerGlySerGlnThrProProAspGluLeuAlaLeuSerGluThrGl 2387
 Db -----GCCTGTGCTCAG----- 4505
 2387 ySerIleSerSerMetProProLeuGluGlyGluLeuGlyAspProAspLeuGluProGl 2407
 Db -----CCCTATGAGGC-----CTAGAGCCGA 4528
 2407 uGlnValGluProGlnProProGlnGlyGlyValAlaAlaProGlySerAspSerGl 2427
 Db AGAGTGT-----GGTATCCTG-----GATGGCTG 4552
 2427 ySerTrpSerThrCysSerGluGluAspSerValValCys---Cys-SerMetSerT 2446
 Db CCCTCAGCGCGCTGTGTACGCGTTCAGAGGCTTCACCTCGGATTTGTTTCATGGCTA 4512
 2446 ySerTrpThrGlyAlaLeuIleThrProCysSerProGluGluGluLysLeuProIleA 2466
 Db CCCTCTGGACA-----TCACCCGCATGCTCTGTGTGCG----- 4644
 2466 snProLeuSerAsnSerLeuLeuArgTyrHisAsnLysValTyrCysThrThrLysAs 2486
 Db ACGTCAACGAGTGTGATGAGGCAGGAGCAACCTCCCCAC-----T 4684
 2486 erAlaSerLeuArgAlaLysLysValThrPheAspArgMetGlnValLeuAspSerTyrT 2506
 Db CTGCATCAATGCGCGCTGTGTCAACACTGACGCGCTCTTCCCGCTGTA----- 4731
 2506 yzAspSerValLeuLysAspIleLysLeuAlaAlaSerLysValThrAlaArgLeuLeuT 2526
 Db TCTGTGCTCCAGAGTTTCTCCAGCATCAGCACATCACTGTGCGCTGCTCG 4786
 2526 hr-----MetGluGluAlaCysGlnLeuThr-----ProProHisSerAlaA 2540

Db 4787 ACCCGGGGCTGAGCAGGGGTGCCAAATCTCCCACTATACCTGAGCACCGGAGCTTCCA 4846
 2540 rgSerLys 2542
 Db 4847 GAGCACGC 4854
 RESULT 9
 BC032614
 LOCUS
 DEFINITION Homo sapiens, Similar to RAB5 interacting protein 3, clone
 IMAGE:5557263, mRNA.
 BC032614
 ACCESSION BC032614.1 GI:21595728
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 2837)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (06-JUN-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 REMARK
 COMMENT Contact: MGC help desk
 Email: cgapps-remail.nih.gov
 Tissue Procurement: ATCC/DCTD/DTTP
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: http://www.nisc.nih.gov/
 Contact: nisc.mgc@nih.gov
 Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantripoop, S., Thomas, P.J., Touchman, J.W.,
 Tsourgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
 Series: IRAK Plate: 69 Row: m Column: 12
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 13376344
 This clone has the following problem: frame shifted.
 FEATURES
 Location/Qualifiers
 1..2837
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5557263"
 /tissue_type="Skin, melanotic melanoma."
 /clone_lib="NIH MGC_72"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 BASE COUNT 581 a 1000 c 805 g 451 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0.0498 Length: 2837
 Score: 175.00 Matches: 208
 Percent Similarity: 31.59% Conservative: 118
 Best Local Similarity: 20.18% Mismatches: 346
 Query Match: 1.09% Indels: 364
 DB: 11 Gaps: 43

Copied from 09540843 on 05/19/2004

Qy 2479 lTyrCysThrThrThrLysSerAlaSerLeuArgAlaLysValThrPheAspArgMe 2499
 Db 1893 GTAC-----AGAGGTGTG---GAGCTGGC 1916
 Qy 2499 tGlnValLeuAspSerTyrThrAspSerValLeuLysAspIleLysLeuAlaSerLy 2519
 Db 1917 GCAGGACAGGGCTCGTACTTGGCAGCCGTGGTGCAGACTACAGGTGTACAGCTGGGA 1976
 Qy 2519 sValThrAlaArgLeuThrMetGluGluAlaCysGln---LeuThrProHisLe 2538
 Db 1977 GATGATGGCGCGCCAGCTCCAGCAGGAGTCTGCAGAGATTCCGACCATGATGAC 2036
 Qy 2538 rAlaArgSerLysTyrGlyPheGlyAlaLysGluValArgSerLeuSerGlyArgAlaVa 2558
 Db 2037 CCAGCTCAAGAGCTACCTGCTGCAGAGCAGCAGCTCAAGCCCTGCTGCAGCCGCT 2096
 Qy 2558 lAsnHisIleLys-----SerValTrpLysAspLeuGlu 2570
 Db 2097 GCATCTCCGAGGAGGCTCGAAGCAATGTAGAGTCTGCTGTGACAAATGTGTCTGAA 2156
 Qy 2570 uAspSerGluThrProThrThrIleMetAlaLysAsnGluValPheCysValAs 2590
 Db 2157 GCCCTGAAGAGCCATCACTCATGCTG-----CATCAGATCCACAGCAGGA 2207
 Qy 2590 pProThrLysGlyGlyLysLysAlaAlaArgLeuIleValTyr-----ProAs 2606
 Db 2208 TGGTTCCTGCAGCAGCTCAAGGAGAACAGTGTAGTCTGCTGCCACCACTGCA 2267
 Qy 2606 pLeuGlyVal-----ArgValCysGluLysMetAlaLeuTyrAspIleThrGlnLysLe 2624
 Db 2268 CCTAGGTGACCAACCAAGCTGCGGAGGTGCCATGATGAGAGATCTCGCAGAGTT 2327
 Qy 2624 uProGlnAlaValMetGlyAlaSerTyrGlyPheGlnTyrSerProAlaGlnArgValG 2644
 Db 2328 C-----ACGAGTGCACAGGCTCTACTCTACCTGAGAGAGATCTC 2369
 Qy 2644 uPheLeuLysAlaTrpAlaGluLysLysAspProMet-----GlyPh 2659
 Db 2370 CATCTCTCAAGACCTGCAAACTCATCTAGACTTCCATGCTCGGCAACCCAGGAA 2429
 Qy 2659 eSerTyrAspThrArgCysPhe-----AspSerThrVa 2670
 Db 2430 GCCTATGCGCGGATGACTTCTGCTGTGCTCATGTATGTGTGCGCCGCGCAGCACT 2489
 Qy 2670 lThrGluArgAspIleArgThrGlu-----GluSerIleTyrArgAlaCysSerLe 2687
 Db 2490 CAGGAGATCTTCTCAATGTGGAGTACATGATGAGCTCATGACCCCTGCTGAGCT 2549
 Qy 2687 uProGluGluAlaHisThrAlaIleHisSerLeuThrGluArgLeuTyrValGlyGlyPr 2707
 Db 2550 GGGGAGGGTCTCTACTA-TCTGACCAACCACT-ACG-----GGGCC 2589
 Qy 2707 oMetPheAsnSerLysGlyGlnThrCysGly-----TyrArgArgCy 2721
 Db 2590 CTGGAGCAGATCAAGAGCTACGACAGATCACCGTGTGCGGAGCTGAGTGTGGAGGT 2649
 Qy 2721 sArg-----AlaSerGlyValLeuThrThrSerMetGlyAsnThrI 2735
 Db 2650 CAGACTCCATCCACCCCTGGAGCGCGGCTACTCTCAACAGGCCCGGCGCTCCCGC 2709
 Qy 2735 eThrCysTyrValLysAlaLeuAlaLysLys 2746
 Db 2710 TCCTCCGTAC---AGGACTTCACTGCTGCTCGT 2740
 RESULT 10
 BC049191
 LOCUS Mus musculus, clone IMAGE:6476607, mRNA.
 ACCESSION BC049191
 VERSION BC049191.1 GI:29165713
 KEYWORDS HTC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 Strausberg, R.
 Direct Submission
 Submitted (21-MAR-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 Contact: MGC help desk
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Contact: nisc.mgc@nih.gov
 Web site: http://www.nisc.nih.gov/
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantropop, S., Thomas, P.J., Touchman, J.W.,
 Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Series: IRAK Plate: 100 Row: 1 Column: 18
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis, genomeScan gene prediction, Similarity but not identity
 to protein
 This clone has the following problem: frame shifted.

FEATURES
 source

1..3929
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:6476607"
 /tissue_type="Colon, normal. 5 month old male mouse."
 /clone_lib="NCI CGAP_C024"
 /lab_host="DH10B"
 /notes="Vector: pCMV-SPORT6"
 BASE COUNT 1293 a 703 c 813 g 1120 t
 ORIGIN

Alignment Scores:
 Pred. No.: 0.104 Length: 3929
 Score: 173.50 Matches: 181
 Percent Similarity: 33.65% Conservative: 105
 Best Local Similarity: 21.29% Mismatches: 304
 Query Match: 1.08% Indels: 261
 DB: 11 Gaps: 39

US-09-980-559-2 (1-3033) x BC049191 (1-3929)

Qy 2182 ProSerHisIleThrAlaGluThrAlaAlaArgLeuAlaArgGlySerProSer 2201
 Db 574 CCTCTGATGAGACCTGTGAGAAAGTGAGAGACAGTAGACTGATATTTAACT 633
 Qy 2202 GluAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThr 2221
 Db 634 GAAGTA--AGTTTCAGTGTGAGCAGCTTGTGCTCTGAGACTAGAGCTGAA---ACACTG 687
 Qy 2222 HisGlyLysAlaTyrAspValAspMetValAspAlaAsnLeuPheMetGlyGlyAspVal 2241
 Db 688 AACAGGAAGCATTTGGAAGTCAAGGAAGGTGAT----- 720

2242 ThrArgIleGluSerGlySerLysValValValLeuAsp-----SerLeuAspProMet 2259
 721 -----GTTCTAGATCCCTGCTTGGACACTTTA 750
 2260 -----ValGluGluArgSerAspLeuGluProSerIleProSer 2272
 751 TCTACTACAGATTTTGTGCTGTGGAAGAA-----GTGGCTCTGCCAAACCCCA 801
 2273 GluTyrMetLeuProLysLysArgPheProProAlaLeuProAlaTrpAlaArgProAsp 2292
 802 AGGCACCTTACTCCAGAA----- 819
 2293 TyrAsnProProLeuValGluSerTrpLysArgPro-----AspTyrGlnProAla 2309
 820 -----CCTGATATAGTACGAGTACAAAGAGCCCTGTTCTCGGGCCCAACCCCA 873
 2310 ThrValAlaGlyCysAlaLeuProProArgLysThrProThrProProArgArg 2329
 874 ACTAATTTCCGGCTCTAGACCCCAACCTCCATCTCGACCTGCACCAACCAAGAAAA 933
 2330 ArgArgThrValGlyLeuSerGluAspSerIleGlyAspAlaLeuGlnGlnLeuAlaIle 2349
 934 AAGAAAAGC-----GAACTGGAGTTTGAAGCTCT 963
 2350 LysSerPheGlyGlnProProPro-----SerGlyAspSerGlyLeuSerThr 2365
 964 AAAACACCTGATCTGGAGTGTACCAAGAAAAATATACATCTGCTTAACTACA 1023
 2366 GlyAlaGly-----AlaAlaAspSerGlySerGln 2375
 1024 AATATGGCTTCAGAGATACAGTTAGAGATTCCTGCTTCCCTCGATTAGCAAGTCT 1083
 2376 ThrProProAspGluLeuAlaLeuSer-----GluThrGlySerIleSerMetProPro 2394
 1084 ACCAGTGGAGATAAATAGTAACCTGCCAGGAAAAATGGAAGCACCTGATGCGAGACA 1143
 2395 LeuGluGlyGluLeuGlyAspProAspLeuGluProGlnValGluProGlnProPro 2414
 1144 GTACAGCGTGAAGTGTGGGCCCTCAG----- 1170
 2415 ProGlnGlyValAlaAlaProGlySerAspSerGlySerTrpSerThrCysSerGlu 2434
 1171 -----AGACCTAGATCCAACTCTGGAGAGAGCTTACTGTAGAGAA 1212
 2435 GluAspAspSerValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuIleThr 2454
 1213 ATTTTAGCCAGTGTATGATTAAGAACCTGGAT-----ACTGGAGAGAAATA- 1260
 2455 ProCysSerProGluGluGluLeuLeuPro-----IleAsnProLeuSerAsnSerLeu 2472
 1261 CCTTTAAGTCTGCTGGAAGAGAGTGTCCAAACAGGTATTAATCTCTCACTCTGCACATC 1320
 2473 LeuArgTyrHisAsnLysValTyrCysThrThrThrLysSerAlaSerLeuAlaGly 2492
 1321 ATGAGA-----AGGACAAAGGATATGTAAGCAATGATGCCACACAGTCAGATGAGAA 1377
 2493 LysValThrPheAspArgMetGlnValLeuAspSerTyrTyrAspSerValLeuLysAsp 2512
 1378 AAGTTACAGTCACAAACACAGATCTGATGGTGAAGGTTTAAACAGAAAAACGACTCAA 1437
 2513 IleLysLeuAlaAlaSerLysValThrAlaArgLeuLeuThrMetGluGluAlaCysGln 2532
 1438 CTAAGAAATTTCTTGGGAAATCAGTGAAGAGCAAGCATCTTGCTGAG----- 1488
 2533 LeuThrProProHisSerAlaArgSerLysTyrGlyPheGlyAlaLysGluValArgSer 2552
 1489 -----GAATATGCT----- 1497
 2553 LeuSerGlyArgAlaValAsnHisIleLysSerValTrpLysAspLeuLeuGluAspSer 2572
 1498 -----GAACTGCTATAAATAAAGTTAAAGCGTA----- 1527
 2573 GluThrProIleProThrThrIleMetAlaLysAsnGluValPheCysValAsp----- 2590

1528 -----AGAGATGAAGTGTCTTCTACTACTGATCAAGAT 1557
 2591 -----proThrLysGlyGlyLys 2596
 1558 GACCTTTCATCAAGTACGATGAAGGATGCCATATACAGACCACTAANA-----TTC 1611
 2597 LysAlaAlaArgLeuIleValTyrProAspLeuGlyValArgValCysGluLysMetAla 2616
 1612 AAGGCAGACATGGTTTCAAGGACCT----- 1638
 2617 LeuTyrAsp-----IleThrGlnLysLeuProGlnAlaValMetGlyAla 2631
 1639 -----TATGATTTTGCATCAAGTGTGCAAGATCTTAGTGTGGAACACATGGAGCT 1695
 2632 SerTyrGlyPheGlnTyrSerProAlaGlnArgVal-----GluPhe 2645
 1696 GTTGGACCATCAAGATTTTCTCAGTGGCCGACTGCTTGTTCAGCTGGACGACCAAC 1755
 2646 LeuLeuLysAlaAlaAlaGluLysLysAspProMetGlyPheSerTyrAspThrArgCys 2665
 1756 ATAGTGAGATATGGCTTTAANAAT-----GCTTTGACTATTTCAACATATG 1806
 2666 PheAspSerThrValThrGluArgAspIleArgThrGluGluSerIleTyrArgAlaCys 2685
 1807 CGAATGAATACATACTCAAGCGCGTGTCACTTCCACATCT----- 1851
 2686 SerLeuProGluGluAlaHisThrAlaIleHisSerLeuThrGluArgLeuTyrValGly 2705
 1852 -----CAGAAAGTCTAAGTTTCATCAAAATCTGACACAGACATGGGGTATGTAGC 1902
 2706 GlyProMetPheAsnSerLysGlyGlnThrCysGlyTyrArgArg-----CysArg 2722
 1903 GGAACGTGACGACCCACATGACAAATGACCGTTTGGCAGCGACCAITTTGCAAA 1962
 2723 AlaSerGly-----ValLeu 2727
 1963 TATAAGGACATACCTGCTGATCTCTTGTATCTTTCATGGTCTAAAAAATCTACTTCTGCTC 2022
 2728 ThrThrSerMetGlyAsnThrIleThrCysTyrValLysAlaLeuAlaCysLysAla 2747
 2023 TCTTCTTCTGATGATAAACAAGTGTGACATTTCTAGAGAGAGATGC----- 2076
 2748 AlaGlyIleAlaProThrMetLeuValCysGlyAspLeuValValIleSerGlu 2767
 2077 -----CTTGTCTGTTTCAACATATAGATTTTGTCACTGCC 2112
 2768 SerGlnGlyThrGluGluAspGluArgAsnLeuArgAlaPheThrGluAlaMetThrArg 2787
 2113 ATCGCTTTCCATCCAGAGATGAC-----AGATA-TTTTCTCAGTGGATCATTTGA 2162
 2788 TyrSerAlaProProGlyAspProProArgProGluTyrAspLeu----- 2802
 2163 TGGAAAGCTCCGCTATGGAACATACCTGACAAAAAGGTGGCTTTTGTGAATGAAGTAGA 2222
 2803 -----GluLeuIleThrSer-----CysSerSerAsnValSerValAlaLe 2816
 2223 TGTCTACACAAAGTGTATCAGCTGCCAATTTTGTGCAATGGCAATATGCAGTAAT 2282
 2816 uGlyPro-----GlnGlyArgArgTyrTyrLeuThrArgAspProThrThrProIleAl 2835
 2283 TGGACATATGATGTTAGTATGTTTCTATGATACAGAG----- 2322
 2835 aArgAlaAlaTrpGluThrValArgHisSerProValAsnSerTrpLeuGly----- 2852
 2323 -CATTTGAATACCATACACAAATACAT-----GTACGATCTACTAGAGCGCAACAA 2375
 2853 -----AsnIleGlnTyrAlaProThrIleTrpAlaArgMetValLeuMetTh 2869
 2376 GGTGGAGAAAAATTTACTGGCATTTGACCTTGGAGAAAAATGAAGATCTGGTAAC 2435
 2869 rHisPhePheSerIleLeuMetAlaGlnAspThrLeuAspGlnAsnLeuAsnPheGluMe 2899

2436	CTCAAAATGACTCAGAAATCAGACTGTATGATGAGTGTGAGACCTTATCACTGTCATGAAGTA	2495
Dbb		
2889	tTyrGlyAlaValTyrSerValSerProLeu-----AspLeuProAl	2903
Qy		
2496	TAAGGGTTATGTCAATAGCAGCGCAGATCAAAAGCAAGTTCAGTCATGACTTTACTTA	2555
Dbb		
2903	alleleGluArgLeuHisGlyLeuAspAlaPheSerLeuHisThrTyrThrProHisG	2923
Qy		
2556	CCTCGTTAGCGGTTCAGAAAGATAAATATGTTTATATCTATCGGAGTACGTAC-----CATGA	2609
Dbb		
2923	uLeuThrArgValAlaSerAlaLeuArg	2932
Qy		
2610	CCTAAGCAAGTTTACTTCTCAGTCAGGAGA	2637
Dbb		

1993	bb	ACATGGGTCCGAGTGGCTTCTTGGAGTGTCTCCGGAGTGGTATCGGTGTCTGTCTGTC	1934
1852	yy	LeuValAspIleLeuAlaGlyTyr-----GlyAlaGlyIleSerGlyAlaLeuValAla	1869
1933	bb	ATCGTGACACTTGTGGTCAGATACCAGGAGGATCCGGA---CAAGGCTCTGTGTGGCT	1877
1870	yy	PhelysIleMetSerGlyGluLysProSerMetGluAspValValAsnLeuLeuProGly	1889
1876	bb	-----TCATGCCCCACCCTCCCATCTCGTCTCT-----GTCTGCCCAGGG	1836
1990	yy	Ile-----LeuSerProGlyAlaLeuValValGlyValIleCysAlaAla	1904
1835	bb	GTCTGGAGGCGGATFCGGCCTCAGCCCTCGGGGATGCGTAATCCCAAGTCTGTCTATGGAA	1776
1905	yy	IleLeuArgArgHisValGlyProGlyGluGlyAlaValAlaGlnTrpMetAsnArgLeuIle	1924
1775	bb	TTCTTC-----AATGATTTCAAGTCGTGGGAGACAAGGTTGGAGTGACATCTCCTT	1722
1925	yy	AlaPheAlaSerArgGlyAsnHisValAlaProThrHisTyrValThrGlu-----	1941
1721	bb	GCT-----GCCGGCGGAAG-GTAGCCCTTCTCAGGCAGGAACGAGAGAGCTTCA	1672
1942	yy	SerAspAlaSerGlnArgValThrGlnLeuLeuGlySerLeuThrIleThrSerLeuLeu	1961
1671	bb	TCACAGCTGGAACTCGGACA-----	1651
1962	yy	ArgArgLeuHisAsnTrpIleThrGluAspCysPro-----IleProCysGly	1977
1650	bb	TACAGGACACAGACACAGCCCTCTGAAGATTGCCCTTGGACTTTCTTCTCTCGGCA	1591
1978	yy	GlySer-----TrpLeuArgAspValTrpAsp-----	1986
1590	bb	GGCAGTTCTCAGGGGTGGCGGATCCCCCTGGATCATCATCCAGGACAAAGCGTCCAAGT	1531
1987	yy	-----TrpValCysThrIleLeuThrAspPheLysAsnTrpLeuThrSerLys	2002
1530	bb	ACAGGTTCTACTGCGCAGACACTGTCCGACTCCT---TCAGTTTGG-----	1486
2003	yy	LeuPheProLysMetProGlyLeuProPheValSerCysGlnLysGlyTyrLysGlyVal	2022
1485	bb	-----AGAGCGCCGGTCTTCAATGGATGGAGTGGCGGTCTTCACAGGGCCACA	1435
2023	yy	TrpAla-----GlyThrGlyIleMetThrThrArgCys	2033
1434	bb	TGGACACAGGCTGGAGGAGTGAAGGCTTCTTGTGCGGGTAAGGCGGTGATCGTCTT	1375
2034	yy	ProCysGlyAlaAsnIleSerGlyAsnValArgLeuGly-----SerMetArgIle	2050
1374	bb	CCCTTCGGCTGCAGA-----GGAAACTGCCGCTTCTCCGGGACTTTTCTCTTCTGTA	1321
2051	yy	ThrGlyProLysThrCysMetAsnIleTrpGlnGlyThrPheProIleAsnCysTyrThr	2070
1320	bb	ACTGCTCACCGGCC-----ACAGGAGCTCTCCT-----	1291
2071	yy	GluGlyGlnCysValProLysProAlaProAsnPheLysValAlaIleTrpArgValAla	2090
1290	bb	-----TCGTGCGCCCCACTCTCTCCGCTCGCTGTCCACAGCGCT-----	1252
2091	yy	AlaSerGluTyrAlaGluValThrGlnHisGlySerTyrHisTyrIleThrGlyLeuThr	2110
1252	bb	-----	1252
2111	yy	ThrAspAsnLeuLysValProCysGlnLeuProSerProGluPhePheSerTrpVal---	2129
1251	bb	-----CCAGCACTCCACCTTCTCCCGGTCCAGTTCCTCTCTCGGT	1210
2130	yy	AspGlyValGlnIleHisArgPheAlaProThrProLysProPhePheArgAspGluVal	2149
1209	bb	GATGGATCTGCC-----AGCGCAGGTCCCATCACCAGGCTGTGGCTGTCCAGCAGATT	1156
2150	yy	SerPheCysValGlyLeuAsnSerPheValValGlySerGlnLeuProCysAppProGlu	2169

Db 1155 TCTGG-TGCACCTCTCTCTCTCTCTTC-----AAGGCCCTTTTGGCCCTCAGCAGAGGAG 1103

Qy 2170 ProAspThrAspValLeuMetSerMetLeuThr----- 2180

Db 1102 TCCAGCTCGCCCAAGCTTTCTCAGTCTCTCTCTCAGCTGTTCCAGGCGGCACTTGAGC 1043

Qy 2181 -----AspProSerHisIleThrAlaGluThrAlaAlaArgArgLeuAlaArgGly 2197

Db 1042 ACAGCCCACTCCACATCCACAGCAGCTTGTCTGCTGGCGTACTGCTGCTGCAGCCGAGC 983

Qy 2198 SerProProSerGluAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuArgAla 2217

Db 982 CGCGCATGCTGCTCTTTCGCGCAGCTCCGACAGCAAGTCTCTCCAGCAGCTTGAGGGCC 923

Qy 2218 -----ThrCysThrThrHisGlyLysAlaTyAspValAsp 2229

Db 922 CTGAAGAGCTGCTGGTTTCCACACCTGTTTGTAAAGTGGTCCCCGCTGAACTCTGT 863

Qy 2230 MetValAspAlaAsnLeuPheMetGlyGlyAspValThrArgIleGlySerGlySerLys 2249

Db 862 AGCCGTGATCGGTTCGCTGATC-----TGTACCAGG--GAGTGTGG----- 821

Qy 2250 valValValLeuAspSerLeuAspProMetValGluGluArgSerAspLeuGluProSer 2269

Db 820 -----GCATGCTCTCCCATG-----GGAAGGCTACAGCCTTGG 788

Qy 2270 IleProSerGluTyMetLeuProLysLysArgPheProProAlaLeuProAlaTTPAla 2289

Db 787 TCAGCATCTCTCGGTCTATCCATCT-----GTCGTGTGA 752

Qy 2290 ArgProAspTyArgProProLeuValGluSerTriPlysArgProAspTyGlnProAla 2309

Db 751 ACCCCC-----TGGACTCAAGCTCAGCAAGCCAGCC 719

Qy 2310 ThrValAlaGlyCysAlaLeu-ProProProArgLysThrProThrPro----- 2325

Db 718 ACATTCTCTGGATGTTGGCTCTCTCTCCAGATCCTCTCCCTCCCTCCATCTGCTTCCCT 659

Qy 2325 ----- 2325

Db 658 TGGCGCCAGTGTTCAGACCCACCTGAGCTTGACGTAGAGAGAGAGATGTTCTGCTTC 599

Qy 2326 -----ProProArgArgArgThrValGlyLeuSerGln 2337

Db 598 AGGACCTGAGCTCTCTGCTGACAGATTTCTCTCGTCTCCACGCTCCGCTTGTGAGG 539

Qy 2337 uAspSerIleGlyAspAlaLeuGlnLeuAlaLysSerPheGlyGln----- 2354

Db 538 GTGTTCTGAGGGCCAGGCTGTGCTCTCCAGTTCTCTCCACGCGAGGGCGTAGAGGTGG 479

Qy 2355 -ProProProSer-----GlyAspSerGlyLeuSerThr----- 2365

Db 478 TCCTCTCTCCATCTCTAGTTGAAACAGCTCTGCGCTCCAGCCCATGTTGGCTCCCGGAA 419

Qy 2366 -----GlyAlaGlyAlaAla-----AspSerGlySerGlnThrPr 2377

Db 418 TCCCCATCCGCGCGCGGTGCAGCTCCGAGTGGCTCGGAGAGCGAGCCCTCTCGCTCG 359

Qy 2377 oProAspGlu-----LeuAlaLeuSerGlu-----ThrGlySerIle 2389

Db 358 TCCCTCTGATCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 299

Qy 2389 eSerSerMetProProLeuGluGlyGluLeuGlyAspProAspLeuGluProGln-- 2408

Db 298 GGGGCCCAACTCTTTCAGGGTCCCGATGGGGAGTCCCGGAGACACAGAGTCACAGTG 239

Qy 2409 -----ValGluProGlnProProGlnGlyGln 2418

Db 238 GAGAGGAGCTGGATGATCTCTGTGAGTGGGATAAGGGCGGAGACCATCCACATCCGG 179

Qy 2418 YValAlaAlaProGly-----SerAspSerGlySerTr 2429

Db 178 GACATCTGCTCCAGGAGGCTGCGAGCTTTCTTCTGAAAGCCCTTCTCTCACATTTGATG 119

Qy 2429 pSerThrCysSerGluGluAspAspSerValValCysCys 2442

Db 118 GTCTCCACAAGTGGGGCTCACTCGGC-----TGTTGC 85

RESULT 12

AK083804

LOCUS

DEFINITION

AK083804 2798 bp mRNA linear HTC 05-DEC-2002

Mus musculus 12 days embryo spinal ganglion cDNA, RIKEN full-length

enriched library, clone:DL130012C14 product:POLYHOMEOTIC 2 homolog

(Mus musculus), full insert sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AK083804.1 GI:26101505

HTC; CAP trapper.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

10349636

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

11042159

3

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,

Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,

Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,

Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,

Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

20530913

11076861

4

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,

Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,

Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,

Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,

Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,

Fleischmann, W., Gaasterland, T., Giusti, C., King, B., Kochiwa, H.,

Kuehl, P., Lewis, S., Matsuo, Y., Nikaudo, I., Pesole, G.,

Quackenbush, J., Schriml, L.M., Stauble, F., Suzuki, R., Tomita, M.,

Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,

Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,

Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,

Fleischer, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,

Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,

Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,

Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,

Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,

Tato, O., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Kohlschki, S.,

Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohlschki, S.

and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection

Nature 409 (6821), 685-690 (2001)

21085660

11217851

5

The FANTOM Consortium and the RIKEN Genome Exploration Research

Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

**JOURNAL
REFERENCE
AUTHORS**

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukaya, S., Furuno, M., Hanegaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Iwata, F., Imotani, K., Iishi, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sugabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Murakami, M. and Havaashizi, A.

TITLE

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suchoi-cho, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216).

URL: <http://genome.gsc.riken.go.jp/>
URL: <http://fantom.gsc.riken.go.jp/>.

FEATURES

```

1. .2798
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:D130012C14"
/db_xref="taxon:10090"
/clone="D130012C14"
/tissue_type="spinal ganglion"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="12 days embryo"
1. .2798
/note="POLYHOMEBOTIC 2 homolog [Mus musculus] (SPTF|Q9QWH1,
evidence: PASTY, 91.81d, 8.4%length, match=219)"

```

BASE COUNT	604 a
ORIGIN	

ORIGIN

Alignment Scores:

Cores:		
	0.214	Length:
	166.50	Matches:
ilarity:	33.74%	Conservative:
similarity:	22.85%	Mismatches:
:	1.03%	Indels:
	11	Gaps:
		2798

US-09-980-559-2 (1-3033) x AK083804 (1-2798)

QY		1996	LysAsnTrpLeuThrSerIlysLeuPheProLysMet-----	2007
			::::: :::	
Db		6	AGAAGCTGGTTTCT-----CTTTTACCATGATGACCATCCTCGAAGACAGCGTAGT	59
QY		2008	ProGlyLeuPro-PheValSerCysGlnIysGlyTyrLysGly-----	2021
Db		60	CCCTCCCTGCCTTGCTGTAAAGCCAGAGACATCACACCGGGGGCTGCTCTGATCTTTGTTTC	119
QY		2022	-----ValTpAlaGlyThrGlyIleMetThrThrAr	2032
Db		120	CCTAGTTTTTCTCTGACTCCGAGCTTGTCGCTGGGGTACTTCTCCCCCTCCCAGGCA	179
QY		2032	gCys-----ProCysGlyAlaAsnIleSerGlyAsnValAr	2044
Db		180	ATSCCTCAGACAGCATATGACTGGACTTCCATGTGCCAGCCAACTGGGACGACCACTCC	239

QY 2346 lnleuAla1lelyS-----SerPheGlyGlnPro-----ProProSerGlyA 2360
 DB 1064 GCTTGGCCCTGAAACCCACACCCAGCAGTAGCCAGCTCTGCTGCGCCACACAGGGCA 1123
 QY 2360 spSer---GlyLeuSerThrGlyAlaGlyAlaAlaSerGlySerGlnThrProProA 2379
 DB 1124 GGACCATGGCTCAGGGTTCCTCCCTGCGGGTCCCAAGCCTAGTGGGACTGCAATGCCCCAG 1183
 QY 2379 spGluLeuAlaLeuSerGluThrGlySerLeuSerMetProProLeuGluGluGluL 2399
 DB 1184 AGACATCAAGGAGAGAT---GGTAACTGCAAC-----ATGAGGGCGCGC 1228
 QY 2399 euGlyAep-----P 2402
 DB 1229 CAGGGCTGGCGGGCAGTTCTGCTGTGCTACTCACCGCTCATTCGACACAGCTTATG 1288
 QY 2402 roaspLeuGluProGluGlnValGluProGlnProPro-----ProGlnGlyV 2419
 DB 1289 CTCACCTGCAGTCACACACAGCTCTCCACAGCGCGCAGGAGCAGCCACAGCC-CCAG 1347
 QY 2419 alAlaAlaProGlySerAspSerGlySerTrpSerThrCysSerGluGluAspSerV 2439
 DB 1348 TTTGTGCGCCACAGACACACAGCCAGCC---ACGACCTGCACCCAGTCCAGTCCCAA 1404
 QY 2439 alValCysCys-SerMetSerTyrSerTrpThrGlyAlaLeuLeuThrProCysSerPro 2458
 DB 1405 CCCAGCTTGCCTCAGTCTCCCAAGC-----CTGGCCCTACAGTCCAGTCCA 1452
 QY 2459 GluGluGluLysLeuProLeuAenProLeuSerAsnSerLeuLeuArgTyrHisAenLys 2478
 DB 1453 GAAGACCATGCTCTGCTCTAGCTCAGTGACACAGGCCCTG-----CCT 1497
 QY 2479 ValTyrCysThrThrThr-LysSerAlaSerLeu---ArgAlaLysLysValThrPheAs 2497
 DB 1498 CTCAGTGTCTACACAGCATGTCACAGCCAGGAGCAGCAGGTGTACCTTCCC 1557
 QY 2497 p-----ArgMetGlnValLeuAspSerTyrTyrAspSerValLeuLys 2511
 DB 1558 ACACAGACACCGGGTCTCAGATG-----GACATCTCTGAG 1593
 QY 2511 sAep-lleLysLeuAlaLeuSerLysValThrAlaArgLeuLeuMetGluGluAlaC 2531
 DB 1594 GGTGGGTCCATCCCTCCCGCAGCGAGGTTCCAGCACACCTC-----AGCT 1638
 QY 2531 ysglnLeuThrProProHisSerAlaArg---SerLysTyrGly-----PheGlyAlaL 2548
 DB 1639 GTGATCTTACA-----GGTGACCGCCGCTTCTCCAGTTACGAGGAGGCTGGTATCA 1692
 QY 2548 ysgluValArgSerLeuSerGlyArgAlaValAenHisLysSerValTrpLysAspL 2568
 DB 1693 AGCCACTGGGGACAGTA---GGGAGGGCCATGCTCAGGTGGTCTGCA-----C 1740
 QY 2568 euLeuGluAepSerGluThrProIleProThrThrleMetAlaLysAenGluValPheC 2588
 DB 1741 TCCTG-----ACTCCA---GGTACATCAGTGTAGGCTAAAGTTGAATA--- 1781
 QY 2588 ysValAepProThrLysGlyLysLysAlaAlaArgLeuLeuValTrpProAspLeuG 2608
 DB 1782 -----TCCCGAGAGAGAGAGGGGTCTCCAGACTTA- 1814
 QY 2608 lyValArgValCysGluLysMetAlaLeuTyrAspThrGlnLysLeuProGlnAlaV 2628
 DB 1815 -----GAGATGACTGAATCTTTCATCTTGGCCGAGCTCCATGGAAACACCGGGTAGTGT 1869
 QY 2628 alMetGlyAlaSerTyrGlyPheGlnTyr-----SerProAlaG 2641
 DB 1870 CCGAGAAGCAGGAGTGGCTCACCAGACGCTCTGTCTGCTCCACCTTTCAGAGCCCGCAGCA 1929
 QY 2641 lnArgValGluPheLeuLeuLysAlaTrpAlaGluLysLys----- 2654
 DB 1930 GTGTGTCCTCCGATGACTGGAGAGAGTGGTGGCCGAGGAGAGTGTGCTGTGCTGCTG 1989
 QY 2655 -----AspProMetGlyPheSerTyrAspThrArgCysPheA 2667

DB 1990 GCCAGCGCCGCTCCCAACACACAGCAAGCATCATCCTGCCATACCGGTGGCTGGCTG 2049
 QY 2667 spSerThrVal 2670
 DB 2050 ACCCAAGAGTC 2060
 RESULT 13
 AK080374/c
 LOCUS
 DEFINITION
 Mus musculus 3 days neonate thymus cDNA, RIKEN full-length enriched library, clone:A630082H08 product:fatty acid synthase, full insert sequence.
 AK080374
 VERSION AK080374.1 GI:26348510
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 REFERENCE
 1 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 99279253
 10349636
 2
 20499374
 11042159
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsumi, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Hashida, K., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861
 4
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Stauber, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Balderelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. P., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohlsuki, S. and Hayashizaki, Y.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)
 21085660
 11217851
 5
 The FANTOM Consortium and the RIKEN Genome Exploration Research

Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

TITLE

JOURNAL
REFERENCE
AUTHORS

6 (bases 1 to 4665)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Saeki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.

Direct Submission

TITLE

JOURNAL

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
Trust/MRC building Addenbrookes Hospital Cambridge) whose
assistance we gratefully acknowledge.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.

FEATURES

source

Location/Qualifiers
1..4665
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:A630082H08"
/db_xref="taxon:10090"
/clone="A630082H08"
/tissue_type="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="3 days neonate"
<1..2939
/note="unnamed protein product; fatty acid synthase
(MGD|MGI:95485, GB|X13135, evidence: BLASTN, 99%,
match=3298)
putative"
/codon_start=3
/protein_id="BAC37895.1"
/db_xref="GI:26348511"
/translation="NVLTRGDLASIRVSWSPKHTQPSSSGAQLCTVYASINFRDIM
LATKSLPDAIPGKWSRDCMLGFEGRDRCRRVNGLVFAGSLATVLLSSDFLWD
VPSWTLDEAASPVVYATYSLVVRGRIQRTGTVLIHSGSGVQQAISIALSLGC
RVFTTVGSAKRAFLQAPQLDDTSFANSRDTSFEQVLLHTGKGVDLVNLSLAE
KLQASVRLAEGKFLIEIGKFDLSNNHPLGMATFLKNVTFHGLLDLALPEANDSWRE
VAALLKAGIDGVKPLCTVFPKAOVEDAFYMAQGHKIGKVLQVVRPEEPAVLPG
AQTPLISAISKTPAKHSYIIIGLGGFLELARWLRLGAQRVLVTSRSGRTGYQ
AKHIREWRGQIQVLTSSVNSVLEGGALAEATKLGVPVGFNLVLRDLWLNQ
TEPLFDQNVKPKYTNLDRATREACPELDYVFVAFSSVSGRGNAGTYNFANSTM
ERICQHRHDLGLAVQWGAIGDVIGVLEAMGTNDVTIGTLPQRISSCMVEYDLFL
NOPHAVLSVLAEGVAGHDGTDQDLKVAHILGRDLNLDLADLADLSDN
LMGVEVQLERHDLVLPREVRQLTLKQESMSKTDSTATTTPKSPSDTSLKON
QLNLTULVNPGEPTLQLNLSVQSSRPLFLVPIEGSTTVFHSLLAKLSVPYIGLQC
TOAAPLDSIINLAAIYIDICIQVQPEGPYRAGYFACVAFECMSQLQAOQGFAPTH
NNLFPDGSITLWYATYQSYRAKMTPCGEAEAEALCFPIKQFLDVHESKVLLELP
LKSIEDRVAASVDLITKSHSLDRRLSFAVSPYHKLRADQVKKAKYHGNVTLIR
AKTGTTGIEDGLADNLSQVCDGKVSVHIIEGDHRTLLEGGSGLESINIITHSLAEP

CDS

VSVREG"
4647..4652
/note="putative"
4665
polyA_site
polyA_site
BASE COUNT 1015 a 1335 c 1310 g 1005 t
ORIGIN
Alignment Scores:
Pred. No.: 0.493 Length: 4665
Score: 166.00 Matches: 298
Percent Similarity: 32.8% Conservative: 156
Best Local Similarity: 21.5% Mismatches: 411
Query Match: 1.03% Indels: 411
DB: 11 Gaps: 73
US-09-980-559-2 (1-3033) x AK080374 (1-4665)
QY 996 AlacysGlyAspIleLeuHisGlyLeuProValSerAlaAArgLeuGlyArgGluValLeu 1015
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3807 AGCTGTGGAGGCCACTTGTGGGGGAATACCATAGTATCAGAGCC----- 3763
QY 1016 LeuGlyProAlaAspGlyTyrThrSerLysGlyTyrSerLeuLeuAlaProIleThrAla 1035
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3762 -----TGAAGCCGGGCGCTG-----ACTGGA 3742
QY 1036 TyrAlaGlnGlnThrArg-----GlyLeuLeuGlyThrIleValValSerMetThrGly 1053
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3741 AGAGACTGAACCGAAGGAGGCTGCTCAGTCAGTCAGTAAATTAATTTGTGGGA---TCAGGA 3685
QY 1054 ArgAspLysThrGluGlnAlaGlyGluIleGlnValLeuSerThrValThrGlnSerPhe 1073
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3684 GAGCATCAAGACGAGCATTTTACAGAGTGTGTGTAATAAATCACACAGTCCACACTTC 3625
QY 1074 LeuGlyThrSerIleSerGlyValLeuThrValTyrHis---GlyAlaGlyAsnLys 1092
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3624 TTCACACTGCAGCTGCAGAAAAGTCAAAACCAAGCATCATTTTCGTTTGGAAACAG 3565
QY 1093 ThrLeuAlaGlySerArgGlyProValThrGlnMetTyrSerSerAlaGluGlyAspLeu 1112
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3564 ACAGTCAGTCCCGGTCGGTCCGAGAGTCAGATGGCCCTGTCATCATGCGGCACGCCGG 3505
QY 1113 ValGlyTyr---ProSerProProGly---ThrLysSerLeuGluProCysThrCysGly 1130
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3504 GGATGCTGGGGACCAAGGGCCAGGGCTACAGTGTCTGTGGGGTCTTGTTCCTCGGGA 3445
QY 1131 AlaValAspLeuTyrLeuValThrArgAsnAlaAspValIleProAlaArgArgGly 1150
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3444 ACCAGCCCTCTCCAT-----ATCCTGCCCTGCAGTGCAGCGGA 3406
QY 1151 Asp-----LysArgGlyAlaLeuLeuSerProArg-----ProLeuSer 1163
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3405 CAGGACACAGACAAAAGGGGCTCCAGCCAGACCCCTGTGTATTTAGTCTAGAACATTTCC 3346
QY 1164 ThrLeuLysGlySerSerGly-GlyProValLeuCysProArgGly-----HisAlaVal 1181
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3345 CAGAAATCTTCCCCACATGCTTCTCCATTTTGGGGGTCTTGAATGAGTCAACCCCTG 3286
QY 1181 lGlyValPheArgAla-----AlaValCy 1189
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3285 GGGGTCTTTCAGCAATGCAATAATATGAGAAATTAATGGAGAAACGACCGCGTGGC 3226
QY 1189 sSerArgGlyValAlaLysSerIleAspPheIleProValGluThrLeuAspIleValTh 1209
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3225 AAATGCCACCGGTAGAAAGGCTCAGTTTGGCTCTGCTCCACTGACTCTTCACAGAC 3166
QY 1209 rArg-----SerProThrPheSer----- 1215
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3165 CAGGGTGTCTCCAAGTCAGTGGCTCTGGGAACCCCACTACGCGGCTGGGAAACACACGCG 3106
QY 1216 ----- 1218
Db 3105 TGATGTTCTGTTGTAGCAGCCACCAGGGGAGCGGTGTGCTGGGCCACTGGGGGAGG 3046

```
Qy 1218 rThrProAlaValPro----- 1224
Db 3045 GGTCCACACAGCAGCAGCCTCCCGAGTTCTGCACTCAGGGTGTGGTGGGATCGGACCA 2986
Qy 1225 -----GlnThrTyrlnValGly----- 1231
Db 2985 TCTCTGTGGCAGAGTGTGGAGCGGTCTCATGGTGTGGCAGGTCTAGCCCTCCCGTA 2926
Qy 1231 rLeuHisAlaProThrGlySerGlyLysSerThrLysValProValAlaAlaG1 1251
Db 2925 CACTCACTGGCTGAGCAGCAGGAGCTATGGATGTTGAT-----GA 2881
Qy 1251 nGlyTyrlsValLeuValLeuAsnProSerValAlaAlaThrLeuGlyPheGlyAlaTy 1271
Db 2880 TGGATTCAGGCCACTCCCTCCAGCAGTGTGGGTGTCACC----- 2838
Qy 1271 rLeuSerLysAlaHisGlyLeuAsnProAsnIleArgThrGlyValArgThrValThrTh 1291
Db 2837 -CTCAATGATGTGCACAGACACCTTCCTCCGTCAACACCTGGGAGAGGTTGTAGTCAAC 2779
Qy 1291 rGlyAlaProIleThrThrTyrlsThrTyrlsPheLeuAlaAspGly--GlyCysAl 1310
Db 2778 C-----CAGTCTCGCCATAGTGTCCCGCTGTCT 2749
Qy 1310 aGlyGlyAlaTyrlsPheLeuAlaAspGlyCysAspGlyCysHisAlaValAspSerThrThrI 1330
Db 2748 TGCC-----ACGACAGTGTACGTT-----GCCATGGT 2719
Qy 1330 eLeuGlyIleGlyThrValLeu----- 1338
Db 2718 ACTTGGCTTGGGCTTTACTATCACTGCTGCCGAGCTTGTGTAGAGGACACAGCGCAG 2659
Qy 1338 pGlnAlaGluThrAlaGlyValArgLeuThrVal---LeuAlaThrAlaThrProProG1 1357
Db 2658 CAAGCTCAGCTCTCGCGGTGCCAGGCTGTGGTCACTTGTAGTATAGGTCCAGCGAGG 2599
Qy 1357 ySerValThrThrProHisProAsnIleGlyGluValAlaLeuGlyGlnGlyGluI1 1377
Db 2598 CAGC-----CACCGATCTTCCAG----- 2580
Qy 1377 eProPheTyrlGlyArgAlaIleProLeuSerTyrlleGlyGlyArgHisIleuLePh 1397
Db 2579 -----GCTCTTTCAGTGGCAGCAGGCGCTCCAGCAGCTTGTGT 2542
Qy 1397 eCysHisSerLysLysCysAspGluLeuAlaAlaAlaLeuArg---GlyMetGlyLe 1416
Db 2541 GTTCCACATCAAGAAACTGCTTTATGAAGAGCATAGGCTTCAGCTCGGCTCGGCTT 2482
Qy 1416 uAsnSer-----ValAlaTyrlsArgGlyLeuAspValSerValI1 1430
Db 2481 CACAGCTTGGGTGTCATCTTTCGCGGTAGTCTGGGTGTAGGCAACACAGTGTGTGTG 2422
Qy 1430 eProThrGlnGlyAspValValValAlaAla-----Th 1441
Db 2421 AGCGTCAAAACAGGAGAGGTGTGTGGTGTGGGCTGGGCTGGGCTGGCCTGCGAGCT 2362
Qy 1441 rAspAlaLeuMetThrGlyTyrlsThrGlyAspPheAspSerValIleAspCysAsnValAl 1461
Db 2361 GGGAGCACATCTCGAAGGCTACACAGC-----TCCAAATGAGTACCCAGCTATGC 2311
Qy 1461 aValThrGlnValValAspPheSerLeuAspProThrPheThrThrIleThrGlnIleVa 1481
Db 2310 GGTAGGTCTCCCTCAGGTGCTGCTTGTATGCAATCTATGTAGTAGGACAGCAGGT-- 2253
Qy 1481 lProGlnAspAlaValSerArgSerGlnArgArgGlyArgThrGlyArgGlyArgLeuG1 1501
Db 2252 -----CGAATGCTATCCAGGGGGGACGCTTGGG 2224
Qy 1501 ylleTyrlArgTyrlVal-----SerThrGlyGluArgAlaSe 1513
Db 2223 TGCACTGCGGCGGTAGGTGGGCACACTGAGCTTGGCAGCAGCACTGTGGAAACAGCGTGG 2164
Qy 1513 rGlyMetPheAspSerVal----- 1519
Db 2163 TGGAAACCTCAATGGGGTGCACAAGCAACAGAGCGCGCTCAGAGCTCTGCACCGAGTTGA 2104
Qy 1520 -----ValLeuCysGluCysTyrlsAspAlaGlyAlaAlaTr 1531
Db 2103 GCTGGGTAGGTAGGACCCCTCAGGGTTCACAGCAGTGT-----GCTCA 2059
Qy 1531 pTyrlGluLeuThrProSerGluThrThrValArgLeuArgAlaTyrlsPheAsnThrProG1 1551
Db 2058 GGTTCAGTGGTCTCTCTTCAGAGAGCTGTC-----ACTCTGG 2020
Qy 1551 yLeuProValCysGln-----AspHisLeuGluPheTr 1562
Db 2019 ACTTGGGGCTGTGCTGTAGTCCAGTGTGAGGAGCAATTCCTGAAGTTTC 1960
Qy 1562 pGluAlaValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThrLysG1 1582
Db 1959 GCAGCTGAGCTGCCGACCTCAGCAT---GGGCAGCACCATCGTGTCTCGTTCCA 1903
Qy 1582 nSerGlyGluAsnPheAlaTyrlsLeuThrAlaTyrlsGlnAlaThrValCysAlaArgAla-L 1602
Db 1902 GGATCTGACGACTTCCACCATCAGCAGTCCAGGCGGAGTCTGCAGGCTGTCTGT 1843
Qy 1602 yAlaProProSerTrpAspValMetTrpLys---CysLeuThrArgLeuLysProT 1621
Db 1842 CCAGGTTAATACCTGCGAGGTCTCGATGCTTAGGATGTGTCTACAGCTTTCACCAT 1783
Qy 1621 hrLeu---ValGlyProThrProLeuLeuTyrlsArgLeuGlySerValThrAsnGluVal- 1639
Db 1782 CCCTCTGGGTGTCCCGTCCCATGGGCCACACACTTCTCTGCCAGCACAAAGTGC 1723
Qy 1640 ---ThrLeuThrHisProValThrLysTyrlleAlaThrCysMetGlnAlaAspLeuG 1658
Db 1722 TCAGGACTGCTGGGGCTGATTCAGAGAGAGTCCAGTACCTCCATGCAGGAGAGATGC 1663
Qy 1658 lu-----ValMetThrSerThrTrpVal---LeuAlaGlyGlyValLeu- 1671
Db 1662 GCTGAGGACGCTACCTCGATGCTGTGTCATGTCGCCCATCGCTTCCAGGACAATGC 1603
Qy 1672 ---AlaAlaValAlaAlaTyrlsCysLeuAla---ThrGlyCysValCysIleIleGlyA 1689
Db 1602 CCAGCTCACCAATGGCAGCCCACTGCAGCAGGCTGGGAGGCCATCGTGTCTGCGCT 1543
Qy 1689 rGluHisIleAsnGlnArgAlaValAlaAlaProAspLysGlu---ValLeuTyrlGluA 1708
Db 1542 GTTCACTATA-----CGCTCTGTAGTGTGGAGCGGTAGTGTAGTTTGGCCAG 1489
Qy 1708 laPheAspGlu-----MetGluGluCysAlaSerArgAlaAlaLeuIleGluGluG 1725
Db 1488 CATTACCACCGCCGACGCTTACAGAGAGAGAGGCCACAAAGTAGTCCAGCTCAGGCGCAGG 1429
Qy 1725 lyGlnArgIleAlaGluMetLeuLysSerLysIleGln-----GlyLeuLeuGlnG 1742
Db 1428 CTTCCCGGGTGGC---CTGTCAAGGTTCAGGTGCTTATTTGGGCTTGTG--- 1376
Qy 1742 lnAlaSerLysGlnAlaGlnAspIleGlnProThrValGlnAlaSerTrpProLysValG 1762
Db 1375 -----ACATCTGTG---AAGAGCT 1360
Qy 1762 luGlnPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrlsLeuAlaGlyL 1782
Db 1359 CTGGGTCTGGTCTCCAGCATG-----GCATCCCTCAAAACCATGCGCCAGGT 1312
Qy 1782 euSerThrLeuProGlyAsnProAla---ValAlaSerMetMetAlaPheSerAlaAlaL 1801
Db 1311 TGAAGACACCCCAACCGGCGCCAGCTTGTGTGGCTTCGGCGATGAGACGACGCGGCC--- 1256
Qy 1801 euThrSerProLeuSer-ThrSerThrThrIleLeuLeuAsnIleLeuGly----- 1817
Db 1255 -----CCCTCCAGTGTAGCTACGCTGCTTGTGTGACACAGCACTTGTAGTCCCTGGC 1204
Qy 1818 -----GlyTrpLeuAlaSer-----GlnIleAlaProPro 1827
```

Db	1203	GTCTCCACTCCCGAATGCTGGCTGGTAGCCGGTGGGATTCAGATCGGGAAGTCA	1144
Qy	1828	AlaGlyAlaThrGlyPheValSerGlyLeuValGlyAlaAlaValGlySerIleGly	1847
Db	1143	GCACAAAGCTCTGGGCTCCGGAAGCAGCAGCCGCGGCGAGCTCCAGGCAAGCCAC	1084
Qy	1848	LeuGlyLysValLeuValAspIle---LeuAlaGlyTyrGlyAlaGlyIleSerGlyAla	1866
Db	1083	CTAGGCC--ACCAAGTGAATGATTAACCTTTATGGGCTGGGCGAAGGCTCTTGAGATGGC	1026
Qy	1867	LeuValAlaPheLysIleMetSerGlyGlyLysProSerMetGluAspValValAsnLeu	1886
Db	1025	-----AGAAATCAGGTGGGCTGAGCCCTGGCGAGCAGCAGCTCAGGCTCCTC	978
Qy	1887	LeuProGlyIleLeuSerProGlyAlaLeuValValGlyValIleCysAlaAlaIleLeu	1906
Db	977	CTCCGTTACCTGGCAGGAGCTTTGCCAATGTGTTCCTCCCTGAGCCATGTAGCGAAGGC	918
Qy	1907	ArgArgHisValGlyProGlyGlyAlaValAlaValGlnTrpMetAsnArgLeuIleAlaPhe	1926
Db	917	ATCTTCCACTGGGCTTGGGAACACTGT-----	888
Qy	1927	AlaSerArgGlyAsnHisValAlaProThrHisTyrValThrGlu-SerAspAlaSerG1	1946
Db	887	GCATTGAGGGCTTCAGACTCCATCAGCAATGCCAGCCTTCAGGAGTCCGCCACCTC	828
Qy	1946	nArgValThrGlnLeuLeuGlySerLeuThrIleThrSerLeuLeuArgLeuHisAs	1966
Db	827	CCGCCAGCTGTCTTGGCTCCTCCTCAAAAGGGGCTCCAGCAGGATCC-ATGGAAGTGA	769
Qy	1966	nTrpIleThrGluAspCysProIleProCysGlyGlySerTrpLeuArgAspValTrpAs	1986
Db	768	CGTTCTTCAAGAGATAG-CCATGCCAGAGGGTGTGTGTAGAAA-----GA	722
Qy	1986	pTrpValCysThrIleLeu-----ThrAspPheLysAsnTrpLeuTh	2000
Db	721	TCAAATTTGCCAATCTCTAAGACGACCATGCTGAGCCAGCAGCCACCTGCGCTGC	662
Qy	2000	rSerLysLeuPheProLysMet-----ProGlyLeuProPhe-----	2012
Db	661	AGCTTCTCTTCTGCGAGTGAGTGTGAGGACAGGTCGAGCCCTTTGCCACCTGTGTG	602
Qy	2013	-----ValSerCysGlnLysGlyTyr-----Ly	2020
Db	601	AACAGTGTCTCAATATGATGTCTCTCGAGTTGGCAAGTGGTGTGTCATCAAGCTGA	542
Qy	2020	sGlyValTrpAlaGlyThrGlyIleMetThrThr-----	2031
Db	541	GGGAACCTGGCTGGA---GGTATGCTGCTTCTCTCGAGGCCACAGTGGTGAAGACG	485
Qy	2032	-----ArgCysProCysGly-----AlaAsnIleSe	2040
Db	484	CGGCAGCCAGACTGAGGGCAATGGAATGGCGCTTTGGCCACACACCTGAACCTGAG	425
Qy	2040	rGlyAsnValArg-----LeuGlySerMetArgIleThrGlyProLysThrCysMe	2057
Db	424	TGGATGAGCAGCGGTCTCCCAAGCTGGATGGCGCCG-----GAACCACTAAGCAGTAG	371
Qy	2057	tAsnIleTrpGlnGlyThrPheProIleAsnCysTyrThrGluGlyGlnCysValProLy	2077
Db	370	TAACAGTGGTAT-----AGAGCGGGCCACAGAGCGCGCTCC	332
Qy	2077	sProAlaProAsnPhelysVal-----AlaIleTrpArgValAlaAlaSerGluTy	2094
Db	331	TCCAGGGTCCAGCTGGAGGGTACATCCAGAGGAAGTCAGATGATAGCAGGACTGAGTG	272
Qy	2094	rAlaGluValThrGlnHisGlySerTyrHisTyrIleThrGlyLeuThrThrAspAsnLe	2114
Db	271	GCCAGGCGCTTCTGAGGAGAACAGCCCATCATCAC-----GCCGGCCACACCTATCCCGCCT	215
Qy	2114	uLysValProCysGlnLeuProSer-----ProGluPhePheSerTrpValAspG1	2131

1963 -----AGTGGG----- 1958
 2253 euAspSerLeuAspProMetValGluGluArgSerAspLeuGluProSerIleProSerG 2273
 1957 -----CCTCATCAATGACGCAAAACCACTGGAGGAGCTGTGCGGCTGG 1912
 2273 luTyrMetLeuProLysLysArgPheProAlaLeuProAlaTatPalaArgProAspT 2293
 1911 AGGAGGCTCCCG-----CCCCCAGCTGCTCAGGTGTGAGCATCAGCAC 1864
 2293 yrAsnProProLeu-----ValGluSerTrpLysArgP 2304
 1863 GTGTACTTGGGCTGCCGAATCTTCGACGACAGATTCGCTGTCTGCTCATGCC 1804
 2304 roAspTyr-----GlnProAla-----ThrValAlaGlyCysAlaLeuProProArgL 2321
 1803 CGAGTGATGACGCGGCTTGAGACACGCTGCGAGGC-----C 1765
 2321 ysThrProThrProPro-----ProArgArgArgArgThrValGlyLeuSerG 2337
 1764 AGACACTGTGTATCCATGATGACAGCAGGAGAGACCAACAGTGAAGC----- 1712
 2337 luAspSerIleGlyAspAlaLeuGlnGlnLeuAlaIleLysSerPheGlyGlnProProp 2357
 1712 ----- 1712
 2357 roSerGlyAspSerGlyLeuSerThrGlyAlaGlyAlaAlaAspSerGlySer----- 2374
 1711 --AGGGCTGCGCGGCTGTAGACAGCGCTGGAGCTGTGTAGCACAGGAGACTTGC 1654
 2375 -----GlnThrProProAspGluLeuAlaLeuSerGluThrGlySerIleSerSerm 2392
 1653 ACTGTAGGAGCAGCACCA-----GCAGCTGGAGATCCACAGCAGATCCGATGATGC 1600
 2392 etProProLeuGluGlyGlu-----LeuGlyAspProAspLeuGluProGluGlnValGluP 2411
 1599 ACGCTCTGCCAGGCGGAAAGGCTTGTGCCCCA-----GCTGTCTCCAGGCGCTGGAACAC 1543
 2411 roGlnProProPro-----GlnGlyValAlaAlaProGlySerAspSerGlySer----- 2428
 1542 CTCAGCGGCGCTCTCTGCCAACTCCCTAGGCGGCCCGCAGGAGTAG-AGTGGCAGCACGG 1484
 2428 ----- 2428
 1483 TGGGGTCCAGGCTGGGCACCTCAGGTACAGTTGTGTGAAGAACAGTGGCTCAGGCC 1424
 2429 -----Trp-----SerT 2431
 1423 CAACAGCATCTGTCTCTCTCCTCACTTGTGCGCGGACACTGGGCTGCCAGTGATCGA 1364
 2431 hrCysSer-----GluGluAspAspSerValValCysCysSerMetSerTyrTrpT 2449
 1363 ACTGCTGCTTCAGGAACACAGACTCCTTGTGTTGTGAC-TGTGGCACCACCCACCCCAAAA 1305
 2449 hr-----GlyAlaLeuLeuThrProCysSerProGluGluGlyLeuProI 2465
 1304 CACTCCCTCTTTCCCGCCACTTCTGCTTCCATCCCTGCTTGGCGAGGAGCTGTACGG 1245
 2465 leAsnProLeuSerAsnSerLeuLeuArgTyrHisAsnLysValTyrCysThrThrL 2485
 1244 AGTCCCGCGCCGCGAGTAGTGTCTGCTTCTGATGTGAGCCGTAGCTAATTCGCCCTG 1185
 2485 ysSerAlaSerLeuArgAlaLysLysValThrPheAspArgMetGlnValLeuAspSerT 2505
 1184 TCATGCGCGCGCAGCCGAGGGAAGATGT----- 1156
 2505 yrTyrAspSerValLeuLysAspIleLysLeuAlaAlaSerLysValThrAlaArgLeuL 2525
 1156 ----- 1156
 2525 euThrMetGluAlaCysGlnLeuThrProProHisSerAlaArgSerLysTyrGlyP 2545
 1155 -----CAGGGGGGCTGTGCTCCTCAGCCCTTCCAGCCCT----- 1123

2545 heGlyAlaLysGluValArgSerLeuSerGlyArgAlaValAsnHisIleLysSerVal- 2564
 1122 -----AGCTTGACTGGAGGGGCTGAGTCCGCTGGTACCTGGGGTTC 1083
 2565 -----TrpLysAspLeuL 2569
 1082 GATGGGCTGCTGACGGGCTGAGGTGGCTGTGCTTACAGGTTCCTCCCTGGAGGCTTCC 1023
 2569 euGluAspSerGluThrProThrIleMetAlaLysAsnGluValPheCysV 2589
 1022 TC-----AACTGCTACAGCC 1008
 2589 alAspProThrLysGlyGlyLysLysAlaAlaArgLeuIleValTyrProAspLeuGlyV 2609
 1007 CCAGCCCTCCGATGGGGTCCAGCTGGCTGCTCTC----- 970
 2609 alArgValCysGluLysMetAlaLeuTyrAspIleThrGlnLysLeuProGlnAlaValM 2629
 969 -----CTGCTGACCTGTGCGGGGCTCTCCCA----- 943
 2629 etGlyAlaSerTyrGlyPheGlnTyrSerProAlaGlnArgValGluPhe----- 2645
 942 --GGGCTCTCTGCTCCATCTCGCTTCTGCTCTCCACTGCTGCTGGGCTGGGGCTCCCC 885
 2646 -----LeuLysAlaTatP-----AlaGluLysLysAspProMetGlyPheSerTyrAsp 2662
 884 ACAGGATGACTGACTTCTTGGAGGCTGAGGCTTGGGCTTGGGAGCCAGCACCGA 825
 2663 ThrArgCysPheAspSerThrValThrGluArgAspIleArgThrGluGluSerIleTyr 2682
 824 CCAAGGACAGCCGACTCACCA-----GGGATCAGAAGTTGTGATTC----- 783
 2683 ArgAlaCysSerLeuProGluGluAlaHisThrAlaIleHisSerLeuThrGluArgLeu 2702
 782 -----TCTGAGCCTAGATCAGGCTCGAGCTTTG----- 753
 2703 TyrValGlyGlyProMetPheAsnSerLysGlyGlnThrCysGlyTyrArgArgCysArg 2722
 752 -----GGGGCCCC-----AGAAATCTGGGACCTCCTGCTGACATCGCTGAAC 708
 2723 AlaSerGlyValLeuThrThrSerMetGlyAsnThrIleThrCysTyrValLysAlaLeu 2742
 707 CAGCCAGGATCTAGGAGCCCGCCAGCCG-----TGGCTCAGGAGTGC 666
 2743 AlaAlaCysLysAlaAlaGlyIleIleAlaProThrMetLeuValCysGlyAspLeu 2762
 665 TGCAGATGCTGGAGCCGCGCTGGCTTGGCTGAGGCTCAGGAGCTGTGGAGCTCATCA 606
 2763 ValValIleSerGluSerGlnGlyThr-----GluGluAspGluArgAsnLeuArgAla 2780
 605 CTGACTTTTCTGCAAGGAGGGGACAGGCTGTACCTTGGGGGCTTTGGGGTGGATGCC 546
 2781 PheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProArgProGluTyr 2800
 545 TTAGATGAGGCTCT-----CCTAGAGCCACCGTCTGGGGCCAGG--- 504
 2801 AspLeuGluLeuThrSerCysSerSerAsnValSerValAlaLeuGlyProGlnGly 2820
 503 -----GCTGCTCGCGGCTGC 489
 2821 Arg-----ArgArgTyrTyrLeuThrArgAspPro-----ThrThrProIleAla 2835
 488 AGGTGCTTTCAGATTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 430
 2836 ArgAlaAlaTatPglu-----ThrValArgHisSerProValAsnSerTrpLeuGly 2852
 429 GCTCCGCGCTGGGTAGGCTGTGACTTCTTGTGCGCAGCCG----- 388
 2853 AsnIleGlnTyrAlaProThr-----IleTrpAlaArgMetValLeuMetThr 2869
 387 -----ATTGATGAGGCGCCCGCAGCAGCGGGGCTCTGG----- 355

Copied from 09540843 on 05/19/2004

2206	QY	rAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThrThrHisGlyLysAlaTy	2226	QY	rAspValAspMetValAspAlaAsnLeuPheMetGlyGlyAspValThrArgIleGluSe	2246	QY	rGlySerLysValValValLeuAspSerLeuAspProMetValGluGluArgSerAspLe	2266	QY	CGGCCAAAGCGCTTCATGCTGGCAAGCTCTTCCCTTTGGTGGCACCGGGCGTGA	2286	QY	uGluProSerIleProSerGluTyMetMetProLysLysArgPheProAlaLeuPr	2306	QY	GGGTCCAAAG-----GGCGGGGGGGGAGCGCGGCGAGATGGAGACCCGAGCAGTGGC	2326	QY	oAla-----TTPAlaArgProAspTyrAsnProProLeuVal	2346	QY	AGCAGCGGGTTACAGAGGAGTGGCGGTGGACGAGG-----CCCTGGCT	2366	QY	lGluSerTrpLysArgProAspTyrGlnProAlaThrValAla-----GlyCysAlaLeuPr	2386	QY	GAGGAGCTGGGAAGCCCCCGGC-----CCGAAGAGGAAGGCTCAGGACACAGCGCTGGCC	2406	QY	oProArgLysThrProThrProPro-----ProArgArgArgArgTh	2426	QY	CCGTTGGGGCGCCCTGGCCCTCCGACCCCGCGAGCGCGGCGCGCGCGCGCGCGC	2446	QY	rValGlyLeuSerGluAspSerIleGlyAspAlaLeuGlnGlnLeuAlaIleLysSerPh	2466	QY	-----CAATGCGCGCTGCCCGCGGCC	2486	QY	eGlyGlnProProProSerGlyAspSerGlyLeuSerThrGlyAlaGlyAlaAlaAspSe	2506	QY	CCGACACCGCCCGCGGTGTGTGCAGTGGG-----ACGGCGCTGGGGCTGGGGGGC	2526	QY	rGlySerGlnThrProProAspGluLeuAlaLeuSerGlu-----Th	2546	QY	GAG-GGTGCCCGCCCGCCGAGCGCTGGCGCTGCCGAGTGCCACCCTGTGCCATGGC	2566	QY	rGlySerIleSerSerMetProProLeuGluGlyGlu-----LeuGlyAs	2586	QY	CGGGCCCCCACCGCCCCCGCCCGCTGCGCGGAGAGCGCCAGCTTTTACAGCTCCCGCC	2606	QY	pProAspLeuGluPro-----GluGlnValGlu-ProGlnProProGlnGlyGlyV	2626	QY	CCCACCGCTACCGCCACCTCGGATCCCTCGAGACCCGAGCGCGCTCCCTAGGCCCGAG	2646	QY	alAlaAlaPro-----GlySerAspSerGlySerTrpSerThr-----	2666	QY	CCCCAGCCCCCAGCGGTGTGCTGGCCAGCGGTGGCATTAGCTTACCGCGCTGGCT	2686	QY	--CysSerGluGluAspAspSerValVal-----CysCysSerM	2706	QY	TCTCTTCTCATATGCAATA-TCAGTCAATTACAGGGGGCGCGGTCTCCAAAATGCAAAA	2726	QY	etSerTyrSer-----TrpThrGlyAlaLeuIleThrProCysSerP	2746	QY	TCGCTTCAGCCCCGAGGGCCCCCAGTATTGGTGGGTGGCCCCCGCATCTGCGGTTTCT	2766	QY	roGluGluGluLysLeuProIleAsnProLeuSerAsnSerLeuLeuArgTyrHisAsnL	2786	QY	TTCCAAAGAAAGGAACCCCAANTCCACAGAAAGAACACACACACACACACACAC	2806	QY	ysValTyrCysThrThrThr	2826	QY	-----	2846	QY	-----	2866	QY	-----	2886	QY	-----	2906	QY	-----	2926	QY	-----	2946	QY	-----	2966	QY	-----	2986	QY	-----	3006	QY	-----	3026	QY	-----	3046	QY	-----	3066	QY	-----	3086	QY	-----	3106	QY	-----	3126	QY	-----	3146	QY	-----	3166	QY	-----	3186	QY	-----	3206	QY	-----	3226	QY	-----	3246	QY	-----	3266	QY	-----	3286	QY	-----	3306	QY	-----	3326	QY	-----	3346	QY	-----	3366	QY	-----	3386	QY	-----	3406	QY	-----	3426	QY	-----	3446	QY	-----	3466	QY	-----	3486	QY	-----	3506	QY	-----	3526	QY	-----	3546	QY	-----	3566	QY	-----	3586	QY	-----	3606	QY	-----	3626	QY	-----	3646	QY	-----	3666	QY	-----	3686	QY	-----	3706	QY	-----	3726	QY	-----	3746	QY	-----	3766	QY	-----	3786	QY	-----	3806	QY	-----	3826	QY	-----	3846	QY	-----	3866	QY	-----	3886	QY	-----	3906	QY	-----	3926	QY	-----	3946	QY	-----	3966	QY	-----	3986	QY	-----	4006	QY	-----	4026	QY	-----	4046	QY	-----	4066	QY	-----	4086	QY	-----	4106	QY	-----	4126	QY	-----	4146	QY	-----	4166	QY	-----	4186	QY	-----	4206	QY	-----	4226	QY	-----	4246	QY	-----	4266	QY	-----	4286	QY	-----	4306	QY	-----	4326	QY	-----	4346	QY	-----	4366	QY	-----	4386	QY	-----	4406	QY	-----	4426	QY	-----	4446	QY	-----	4466	QY	-----	4486	QY	-----	4506	QY	-----	4526	QY	-----	4546	QY	-----	4566	QY	-----	4586	QY	-----	4606	QY	-----	4626	QY	-----	4646	QY	-----	4666	QY	-----	4686	QY	-----	4706	QY	-----	4726	QY	-----	4746	QY	-----	4766	QY	-----	4786	QY	-----	4806	QY	-----	4826	QY	-----	4846	QY	-----	4866	QY	-----	4886	QY	-----	4906	QY	-----	4926	QY	-----	4946	QY	-----	4966	QY	-----	4986	QY	-----	5006	QY	-----	5026	QY	-----	5046	QY	-----	5066	QY	-----	5086	QY	-----	5106	QY	-----	5126	QY	-----	5146	QY	-----	5166	QY	-----	5186	QY	-----	52
------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	---	------	----	--	------	----	--	------	----	--	------	----	--	------	----	---	------	----	--	------	----	--	------	----	----------------------------	------	----	--	------	----	---	------	----	---	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	--	------	----	---	------	----	--	------	----	---	------	----	---	------	----	--	------	----	--	------	----	--	------	----	----------------------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	------	----	-------	----

Search completed: February 12, 2004, 09:33:14
Job time : 12477 secs

